Responsible investing

Borgers, A.C.T.

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Responsible Investing:
New Insights into Performance and Tastes

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof.dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op vrijdag 12 september 2014 om 10.15 uur door

Adrianus Carolus Thomas Borgers

geboren op 29 maart 1987 te Veldhoven

Promotiecommissie:

Promotores:	prof. dr. C.G. Koedijk
	prof. dr. J.R. ter Horst
Copromotor:	dr. J.M.M. Derwall
Overige leden:	prof. dr. D. Brounen
	prof. dr. F.A. de Roon
	prof. dr. R.M.M.J. Bauer
	prof. dr. S. Pouget
	dr. P.C. de Goeij

To my parents with love and gratitude

Acknowledgements

Although my name is on the cover, this dissertation would not have been written without the support and help of many others. Therefore, I will start by thanking my supervisors, Kees Koedijk, Jenke ter Horst and Jeroen Derwall for believing in me since I came to them at the end of the first year of the Research Master. Kees has allowed me to use his extensive network throughout my PhD time, enabling me to meet and work with many interesting and influential people. Not only did I have the dean of Tisem and later TIAS in my team, also the vice-dean Jenke ter Horst was willing to work with me. Jenke always gave very detailed comments on my writing and we had valuable discussions about research that after some time diverged to talking about our families. Working with Kees and Jenke gave me some insights into how big institutions and scientific journals work, which I find very intriguing. One consequence was that both of them had limited time available but this resulted in a valuable lesson: how to hold short meetings (something I struggled with). Then there is Jeroen Derwall, my daily supervisor and dear friend whom I can always ask for advice. We have maintained a constant online connection for about 2.5 of my 3.5 year PhD period. Only after Eliza said to me (at their wedding) “So you are the guy who is always asking questions on Skype?” I reduced my contact moments to twice a week instead of twice a day. I am very glad that we continue working together in Maastricht.

I would also like to thank the members in my dissertation committee, Rob Bauer, Dirk Brounen, Peter de Goeij, Sebastien Pouget and Frans de Roon for taking the time to carefully read my work and give valuable feedback. Sebastien, we met on multiple conferences and I hope many others will follow.

Due to the previously mentioned network effect, I was able to work with very good researchers. Rachel Pownall, who gave me so much freedom to explore my ideas. I am really

exited that we keep on working together on the project that we undertake now together with Louis Raes. My final Chapter is work together with Jeroen, Nadja Guenster and Paulo Rodrigues, this hodgepodge of different skills and personalities works great. Paulo, a special word of thanks to you for teaching me some valuable programming skills.

The group of people who made my time in Tilburg legendary are the members of our beloved lunch group, nested in this group is the soup club and the staircase workout crew. Aida, Amparo, Hettie, Jochem, Jongyook, Justin, Marc, Mitzi, Niels, Louis, Rob, Roxana, Sybren, Thomas, and Tim. Everyday around noon we spent about 45 minutes together that I truly enjoyed. We had entertaining discussions with Louis, the vegetables of Thomas, the Latin parties from Aida and Marc, Nielsdag, the dancing of Jongyook, and the treats from Hettie's bakery. I am sure that many more memorable times will follow even though we live far apart. Thomas, I have learned a lot from you and I appreciate that we managed to create a well balanced working atmosphere in the office.

Besides the lunch group which was a melting pot of cultures and departments, I am also indebted to the Finance department. I was not always visible since my office was at the fourth floor. Nevertheless, everybody was always willing to discuss research or update me with the news of the department. The seminars in Tilburg are great and I hope that the department will continue to do so well. A special word of thanks for Jos Grazell and Peter de Goeij since I really enjoyed teaching in your courses, and Luc Renneboog, for your trust in me and supporting me to enter this prestigious program. I will also miss the coffee break discussions with Vincent and Martijn.

As said, my office was at the fourth floor of the Koopmans building, with the Tilburg Sustainability Center. This group works on sustainability issues from different angles. It has been interesting to learn about, economic models, law, psychology, and so on. TSC, thank you for broadening my scope.

I spent my last three months in Tilburg with TIAS. The people there made me feel at home from the beginning. The limited distraction and dynamic atmosphere helped me to implement all comments from the dissertation committee.

My friends from Veldhoven provided me with valuable distractions in multiple forms. Informal dinners with “the couples”, playing squash with Jeroen, Tijs and Thomas, motorbike riding with the notorious Motorwhiskygenootschap and the Rietbeek M.C., parties and get-togethers with the guys. Thank you all for being such amazing company. Jeroen, I have really appreciated all our meals and drinks with good talks.

Great support came from my family during my Research Master and PhD period. They have always been there to help with more practical matters or listen to my frustrations about (economic) inefficiencies. Ad, Elly en Familie Kappen, bedankt voor alle gezelligheid. Mariolein en Rene, Charlotte en Stefan, Joris en Danielle, ik geniet altijd van onze familie-uitjes en ik hoop dat we nog veel weekendjes weg kunnen plannen. Papa en mama, bedankt dat jullie altijd voor me klaar staan. Dankzij jullie heb ik de juiste mentaliteit gekregen om dit proefschrift succesvol af te ronden. Als laatste wil ik mijn grote liefde, Loes, bedanken. Ook al was het soms moeilijk voor jou omdat ik vaak in gedachten nog met mijn onderzoek bezig was, je zorgde altijd dat ik fijn thuis kon komen. Ik hoop nog voor altijd bij jou thuis te komen.

Table of contents

Chapter 1

1. Introduction

"At Parnassus Investments, we follow a responsible investment approach to understand the full impact of a company. We carefully consider a company's environmental, social and governance (ESG) factors. By incorporating ESG factors into our fundamental investment process, we often identify risks and opportunities that the market may have ignored, and identify responsible companies."[1]

This quote from an investment fund that explicitly practices socially responsible investment, or SRI, is an example of how investors nowadays are embracing this investment approach. SRI emerged from a faith based investment practice to evolve into a broader consideration of non-financial indicators that determine a firms' impact on society and the environment (Renneboog et al. 2008). The remains of this faith-based approach are still present in some SRI labeled investment funds in the form of exclusionary investment screens.[2] These screens determine the investable universe, companies that act in contradiction with the investor's norms, values, beliefs, or tastes are not considered as an investment:

"Parnassus Investments applies exclusionary screens to our investment universe. The spirit of these exclusionary screens, as described in the Funds' prospectus, is to avoid investment in companies with negative impacts that outweigh any potential benefits from their business activities.

We do not invest in companies that derive significant revenue from the following activities:

[1] www.parnassus.com, retrieved early 2014.

[2] Also refered to as "negative screens".

. *Manufacture of alcohol products*
. *Manufacture of tobacco products*
. *Direct involvement with gambling*
. *Manufacture of weapons*
. *Generation of electricity from nuclear power*
. *Business involvement with Sudan*

…".

Besides this fairly simplistic form of SRI, investors use information about corporate social responsibility (or CSR) to make their investment decisions. Investors refer to this type of information as environmental, social, and governance (or ESG) information. One way to use this information is by applying best-in-class screening, meaning that investors consider only the companies with the best CSR practices (as measured by ESG information) compared to firms in their industry.[3] Investment companies generally do not commit to this form of screening, instead they state the use of ESG information where relevant:

"Parnassus Investments uses qualitative analysis of environmental, social and governance (ESG) factors to help determine if a company is appropriate for investment. ESG analysis can provide insight into a company's management, culture and competitiveness, as well as its impact on society. We believe that integrating ESG analysis into our investment process can help minimize investment risk and improve returns.".

In addition to using ESG information when investing and applying exclusionary investment screens, investors also engage with companies in attempts to change their behavior (e.g. Dimson et al. 2012). Consider environmental externalities; engagement with companies to reduce these externalities can lead to a change in the company's behavior, e.g. a reduction in externalities (Chava 2013). Successful engagement can even be profitable for the

[3] Also refered to as "postive screening".

firms as well as its investors (Dimson et al. 2012). In contradiction, the screens in the example above are very stringent; Phillip Morris' main products are cigarettes, therefore this company will not be considered as an investment as long as the tobacco exclusionary screen is applied. Throughout this dissertation, I will focus on SRI screens.

1.1 Effects on performance: tastes vs. mispricing

Before I present the research of this dissertation, I will continue with a literature review on the current state of the SRI literature with a focus on the effects of SRI screening on returns of equities. Subsequently I will use this review as background for the chapters in this thesis. The majority of papers that look into the effects of SRI on equity returns hypothesize that exclusionary SRI strategies should have a negative effect on SRI portfolio performance whereas the use of ESG information might have a positive effect through increased understanding of the firm. Below I will explain these hypotheses and their foundations as well as the hypothesized effects on individual equities.

Effects of tastes: According to the widely used Markowitz portfolio theory, SRI (in the form of restrictions on investment opportunities) should be costly because limiting the investment space can't lead to better outcomes in equilibrium. The negative effects on an investment portfolio can come from a loss of diversification or missed investment opportunities. Chapter 4 directly investigates tastes for SRI screens of Dutch pension fund beneficiaries. If they get positive utility from SRI they might even be willing to accept the hypothesized lower return resulting from the investment restrictions.

Besides affecting the portfolios of investors that use SRI screens, tastes for or against assets has the potential to influence the portfolios of all investors. In the following I will provide an example applying the capital asset pricing model (CAPM), for simplicity assume there is one group of investors with similar tastes, one group without tastes, and rational asset prices. What will happen is that the group without tastes holds the optimal unconditional

mean variance efficient portfolio (T), the investors without tastes hold the *conditionally* mean variance efficient portfolio (D). The market portfolio (M) is the positively weighted average of T and D on the hyperbola connecting them, see Figure 1.1. As a result, M is generally not mean variance efficient and we do not get CAPM pricing. However, since all assets need to be held, we will get to a situation where the investors without tastes offset the holdings of the investors with exogenous tastes. Since the investors without tastes should deviate from the optimal portfolio, they want to be compensated in the form of higher expected returns (or lower prices) for the stocks they have to overweight. This effect can work in the opposite direction if a stock is very popular, driving up the price.

The logic, even when using other asset pricing models, is that tastes can shift prices of securities (that are subjective to tastes) away from their "true" value as if tastes were not present. This happens if similar tastes influence the portfolios of a significant number of investors. Consequently the expected returns of those stocks can be abnormally high or low and this can be tested using measures of Jensen's alpha (e.g. Fama and French 2007, Hong and Kacperczyck 2009).[4] Applying this theory to the effects of exclusionary SRI screens on individual securities, we should see that the excluded controversial stocks trade at lower prices and have higher expected returns, given that similar screens are applied by a significant number of investors. Similarly, best-in-class screens can lead to the overweighting of stocks with good CSR profiles leading to higher prices and lower expected returns for these stocks.

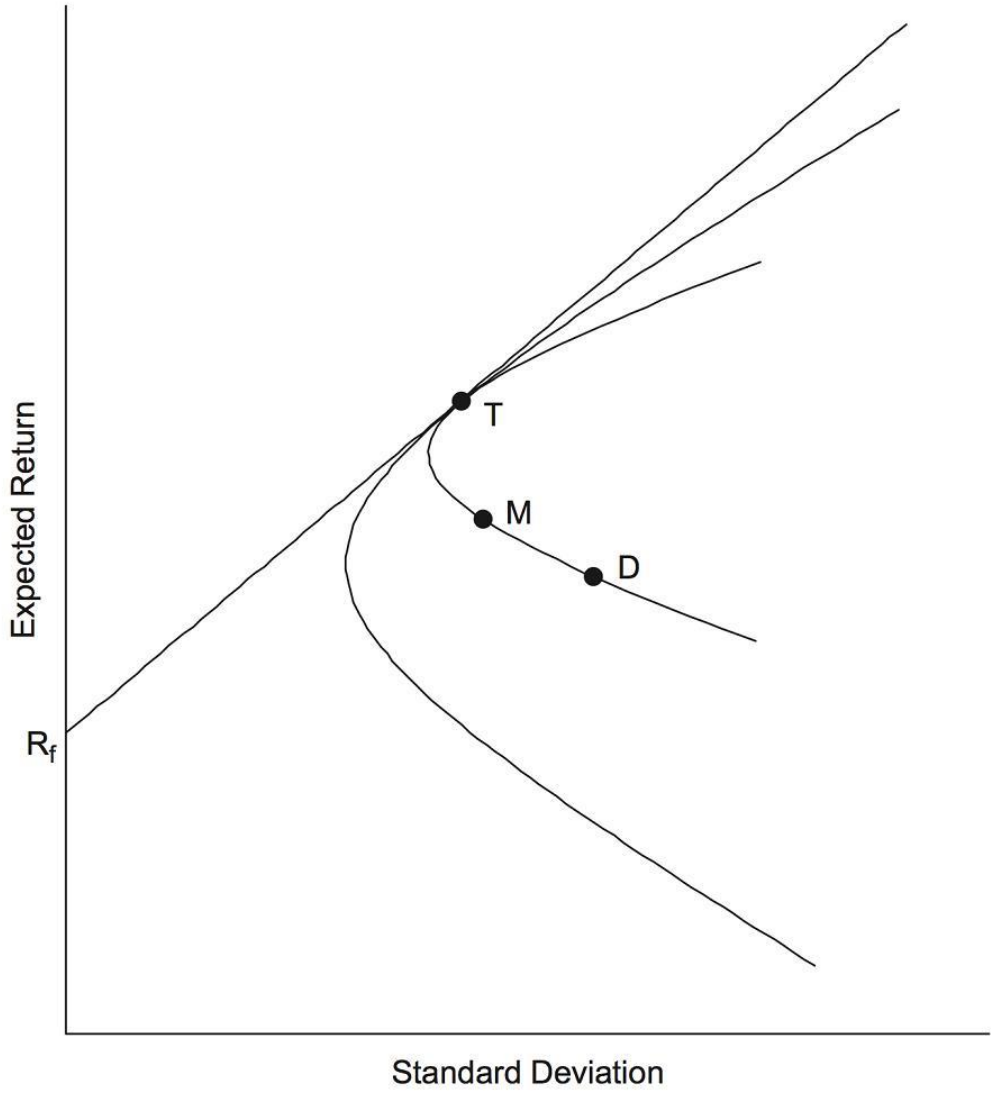
Alternative explanations for asset prices deviating from the CAPM also exist. Merton (1987) argues that deviations can be traced back to neglected firms having a higher perceived risk due to information asymmetry between investors. Using this theory, one could argue that investors disagree about the payoff structure instead of having tastes for or against assets. As

[4] See Fama and French (2007) for a more general and formal proof as well as specific situations in which (I)CAPM pricing holds.

a result, the average belief about the payoff structure *can* deviate from the "true" or "rational" one leading to irrational prices (Fama and French 2007). These alternative arguments also support the idea that an SRI portfolio excluding certain assets from the investment space performs similar to a conventional portfolio at best.

Figure 1.1 The CAPM with segmented investors

This figure shows investment opportunities for investors under the Capital Asset Pricing Model. T is the unrestricted mean variance efficient tangency portfolio. D is the conditional mean variance efficient portfolio for investors with a restricted universe. M is the market portfolio which is the weighted average of T and D. For more information see Fama and French (2007).



Effects of mispricing: On the other hand, the use of value relevant ESG information used in SRI might be overlooked by conventional investors. Meaning that the conventional investors do observe the costs that are immediately expensed through the income statement but they undervalue the benefits of ESG practices that are often intangible, difficult to quantify, and materialize in the long term similarly to R&D investments (Lev et al. 2005, Derwall et al. 2011). The idea of undervalued ESG practices can be justified by both instrumental stakeholder theory (e.g., Cornell and Shapiro 1987, Zingales 2000) and the resource-based view of the firm (e.g., Wernerfelt 1984, Barney 1991, Hart 1995, Russo and Fouts 1997).[5]

Both theories predict that ESG practices can be value relevant. Therefore, informed investors who trade based on this information can reap the benefits of underpriced ESG practices. This can be done by identifying underpriced CSR activities of firms, make the investment in the firm before the market recognizes the underpricing, and finally hold on to the investment until the market corrects the price. Particular moments when such corrections take place are earnings announcements since this is when the market should recognize that earnings are higher than expected (Core, Guay, and Rusticus 2006, Edmans 2011).

In line with the view that ESG practices *can* lead to better performance, the empirical literature that tests if doing well results in doing good, finds mixed results, see McWilliams and Siegel (2000) for a review. However, these papers cannot conclude that a causal relationship exists, e.g. firms with higher (anticipated) accounting profits have more possibilities to invest in CSR. Recent work of Jiao (2010) and Deng, Kang, and Low (2013) use instrumental variables and find a causal relationship of corporate social responsibility on

[5] In instrumental stakeholder theory, the firm is a nexus of contracts and has the function of a middleman between the consumer and the suppliers of inputs. Therefore the claimants of a firm are all parties with explicit or implicit contracts with the firm. The highest firm value is reached when those contracts are managed optimally. The resource-based view of the firm argues that CSR practices that are pro-active and value relevant can lead to a competitive advantage just as any other investment.

accounting profits. While Cheng, Hong, and Shue (2014) find evidence that CSR is at least partially explained as managerial pet projects using two quasi-experiments to control for causality. This evidence shows that the relationship between CSR and earnings is complex.

Empirical evidence of socially responsible investment returns

In the following I will describe studies on the returns of investments with an SRI dimension. Using the two effects described above I will summarize and comment on the findings of studies into SRI. Looking only at the effects of tastes one would predict that SRI yield lower returns due to increased investment restrictions while the effects of mispricing predict better performance conditional on CSR investments by companies being undervalued. I will start with an evaluation of hypothetical stock portfolios that enable me to test the hypotheses separately. Subsequently SRI mutual fund return studies will be evaluated.

Effects of tastes on investment portfolios

Putting extra constraints on a portfolio should lead to lower expected returns of this portfolio. A common constraint in SRI is excluding stocks that go against social norms. If enough investors neglect these stocks, the prices go down and the risk-adjusted returns go up as the investors that do invest in these stocks have a smaller investor base for risk sharing and thus require an extra risk premium (Heinkel, Kraus, and Zechner 2001). This indirect way of testing the effects of tastes on performance is referred to as the "shunned stock"- or "neglected stock" hypothesis (see e.g. Statman and Glushkov 2009). Common examples of such shunned assets are stocks of companies that earn revenues from the tobacco, alcohol, gambling, and weapons industries (e.g., Hong and Kacperczyk 2009), or stocks of heavily polluting firms (Heinkel, Kraus, and Zechner 2001)[6].

[6] Other related theoretical studies include Angel and Rivoli (1997), Fama and French (2007), and Gollier and

Empirical evidence teaches us that socially controversial stocks have either higher expected returns as implied by stock prices or earned higher realized returns than socially acceptable stocks (e.g., Fabozzi, Ma, and Oliphant (2008), Hong and Kacperczyk (2009), El Ghoul, Guedhami, Kwok, and Mishra (2011), Derwall, Koedijk, and Ter Horst (2011), Chava (2013)). In Table 1.1 I provide an overview of studies into controversial investment returns.

However, for social norms to affect asset prices the requirement is that enough assets under management are influenced by these social norms (Heinkel et al. 2001). Testing this requirement, Hong and Kacperczyk (2009) show that institutional investors whose investments are influenced by social norms (including pension funds, banks, university endowments, employee ownership plans and others) hold fewer assets in sin stocks. In line with the hypothesis that sin stocks are neglected by influential investors they find that compared to a sample of equivalent firms without activities that go against social norms, sin companies are financed more with long-term debt and are covered by fewer analysts. Finally, the result for equity owners has been a four-factor risk-adjusted outperformance of 26 basis points per month for a portfolio of sin stocks minus comparables over the period 1956-2006.

Other papers on controversial investments have found similar return effects. Kempf and Osthoff (2007), Statman and Glushov (2009), and Derwall et al. (2011) all report positive abnormal returns for a portfolio of controversial stocks based on KLD STATS' controversial business indicators. Papers that do not rely on a social ratings agency used industry selection measures like Hong and Kacperczyk (2009). They also find statistically significant positive risk adjusted returns (Table 1.1). Studies specifically into international controversial stocks find that a portfolio of European sin stocks minus a portfolio without sin stocks yields a risk adjusted 4% annually over the period 1975-2006 (Salaber 2007). And broader domestic

---

Pouget (2012).

portfolios of controversial stocks, producing an average 19% annual return, outperform the domestic market indexes in a study that involves 21 countries (Fabozzi et al. 2008).

However, more recent papers find that the shunned stock effect is in fact country specific. Durand et al. (2012) report *underperformance* for sin industries in 7 Pacific Basin markets. They use the cultural Hofstede dimension of individualism to support their argument that in more (less) individualistic countries, investors herd away from (towards) sin stocks arguably because the more individualistic investors believe they are accountable for their own actions. Fauver and McDonald IV (2014) separate the G20 members into "sin" and "non-sin" countries by measuring a social norm against sin products. This measure uses the world values survey, consumption data of sin products, and legal restrictions towards sin products. Sin stocks had abnormal returns of 1-2% annually over the 1995 to 2009 period only in countries that view these products as sinful compared to other countries. In addition, they find that the sin stock premium only stays in existence over time when there are barriers to capital and language. These barriers ensure that not enough arbitrage capital enters those markets to offset the shunned stock effect.

A recent addition to the empirical literature on controversial investment returns is on environmental concerns, Chava (2013) reports that the shunned stock effect can also be applied to stocks with environmental concerns like pollution or extreme greenhouse gas emissions. He uses the internal cost of capital method of Gebhardt, Lee, and Swaminathan (2001) to estimate the cost of equity capital and the cost of debt capital. Both measures are affected by the environmental profile of the firm. Firms with more environmental concerns have higher cost of equity and debt capital while firms with more strength indicators only have cheaper debt financing. The results support the story that the financial market is able, at least to some extent, to let firms pay for their environmental externalities. The difference with

the sin industries is that firms can change their environmental profile while staying in the same line of business.

In contrast to the previously discussed studies that look at neglect of stocks, Galema et al. (2008) directly test the hypothesis that CSR of firms affect the demand for their stocks leading to higher prices for responsible firms. Therefore, the book to market value is affected by CSR scores which lead to the HML factor (based on book to market values) included in most SRI performance studies picking up at least part of the return effects that stem from demand driven effects. El Ghoul et al. (2011) also test the hypothesis that market values are affected by demand for social responsibility. They find that companies with relatively high STATS scores (especially Employee relations, Environmental policies, and Product safety) have a lower ex ante cost of equity capital (in other words, these stocks trade at higher prices). In line with the studies on controversial stocks, they report a higher cost of equity financing for Tobacco and Nuclear energy firms.

Table 1.1 Studies on the performance of controversial stocks

This table is taken from Derwall et al. (2011), slightly modified and extended with more recent studies. "Sin" is an acronym for Alcohol, Gambling, and Tobacco.

Study	Region and period	Sin	Weapons	Nucl.	Biotech	Adult	Env. Con.	Alpha
Hong and Kacperczyk (2009)	US 1926-2006	X	X					Positive
Kempf and Osthoff (2007)	US 1991-2004	X	X	X				Positive (NS)
Statman and Glushkov (2009)	US 1991-2007	X	X	X				Positive (NS)
Derwall et al. (2011)	US 1992-2008	X	X	X				Positive
Salaber (2007)	Europe 1975-2006	X						Positive
Fabozzi et al. (2008)	21 Countries 1970-2007	X	X	X	X	X		Positive
Visaltanachoti et al. (2009)	China 1975-2006	X						Positive
Liston and Soydemir (2010)	US 2000-2007	X						Positive
Fauver and McDonald IV (2014)	19 Countries 1995-2009	X						Positive*
Kim and Venkatachalam (2011)	US 1988-2006	X				X		Positive
Lobe and Walkshausl (2011)	51 countries 1995-2007	X	X	X	X	X		Neutral
Durand et al. (2012)	Pacific Basin 1990-2009	X	X					Negative
Chava (2013)	US 1992-2007						X	Positive

* Only in countries with a social norm against sin products.

The effects of tastes described above are at the *stock* level. Another strand of literature looks into the effects of screens on *portfolio* performance. Hertzel et al. (2012) show that the effects on performance are limited, only when the restrictions on the investable universe become very severe (at least a 30% reduction) do they find significantly negative effects on performance in some cases depending on the screening criteria. Boudt et al. (2013) show that SRI screens theoretically lead to the investor being unable to reach all parts of the unconstrained efficient frontier. However, in their empirical tests the results are generally not statistically significant. Another optimization test using mutual fund returns concludes that investing only in SRI mutual funds can have a negative effect on performance depending on the investment beliefs of the investor (Geczy et al. 2006).[7]

Summarizing, according to theory excluding stocks from the investable universe should lead to negative effects on the investment portfolio and if enough investors exclude certain stocks it can have effects on the cost of equity capital for the excluded companies. Empirical studies have found that indeed stock boycotts lead to firms having a higher cost of equity capital and consequently produce higher risk adjusted returns. The effects differ between regions and the type of business the company operates in as believes about *what* is controversial differs between regions, time periods, and other factors. The effects of tastes also work in the opposite direction leading to a lower cost of equity capital for firms with good CSR profiles that manifest itself through higher ESG indicators. This lower cost of equity capital should lead to these firms having lower expected stock returns. Although there is evidence on individual stocks and stock classes (e.g. sin stocks) there is little evidence on the effects of tastes on investment portfolios. In Chapter 3 I will investigate to what extent

[7] For an investor who believes in the CAPM and not in fund manager skill the cost of the SRI constraint is just a few basis points per month as measured in certainty equivalent loss. While an investor who believes in a multifactor asset-pricing model and allows for manager skill loses at least 30 basis points per month. These findings can be explained by the fact that less stocks are needed to mimic the optimal CAPM portfolio compared to a mean variance optimal multifactor asset pricing model portfolio.

SRI related tastes affect the portfolios of *all* US equity mutual funds and subsequently report on the performance effects of tastes as measured by exposures to for instance firms with very good/bad CSR profiles.

Effects of mispricing on investment portfolios

So far I have discussed the effects of tastes on returns of individual companies as well as the effects on investment portfolios all of which predict negative effects for SRI portfolios. However, integrating ESG information in the investment process can enhance returns through a better understanding of firm performance. If ESG information is relevant for firm valuation and the market ignores or doesn't correctly price this information, the investors that do take this information into account can outperform the market. In line with this hypothesis researchers have reported positive risk adjusted returns for portfolios of equities selected using ESG information from different data sources. The first study forming portfolios on environmental (E) information finds that an industry adjusted portfolio long "eco-efficient" firms and short less "eco-efficient" firms by 4.15% a year from 1995 to 2003. Studies that followed used different datasets and a different time span enabling them to form portfolios based on other characteristics and to check the robustness of the Derwall et al (2005) study. Kempf and Osthoff (2007) also report performance differences between portfolios of stocks with high- minus low ESG scores formed using the KLD STATS data from 1992 to 2004. They report statistically significant risk-adjusted outperformance for the portfolios formed using Community, Employee relations, and Diversity (S) indicators separately. Furthermore, they find that more restrictive best-in-class screening (adjusting the firm specific scores for the respective industries) leads to the highest alpha. Their combination portfolios that apply multiple screens and vary in intensity (Top vs. Bottom 5%, 10%, and 25%) yield abnormal returns ranging from 3.6% to 8.7% annually. Statman and Glushkov (2009) use similar data,

however they exclude all firms that do not receive a single indicator in the KLD STATS database because of the possibility that KLD did not consider these firms in their rankings. From their paper I want to highlight that again only particular screens lead to abnormal performance as well as the overall strategy (long (short) in stocks with high (low) aggregated ratings) yielding an annual value (equal) weighted four-factor alpha of 5.0% (5.5%).

These performance studies base their hypothesis for outperformance on mispricing. Nevertheless, their focus on portfolio returns is merely an indirect way to test for mispriced information, demand effects also have the potential to temporarily drive up stock prices leading to abnormally high returns. A more recent study by Edmans (2011) does perform direct mispricing tests using Forbes' list of "Best companies to work for" between 1984 and 2009. He finds that stocks of the best companies to work for have higher risk adjusted stock returns (a value weighted portfolio earns a 3.5% four factor alpha annually), higher earnings announcement returns, and higher analyst forecast errors than other stocks.

In this dissertation I focus on the effects of Environmental and Social performance indicators, though a third dimension often considered in SRI is Governance, hence the ESG acronym. Gompers et al. (2001) formed portfolios based on their G-index (higher score means lower shareholder rights), created using governance indicators. These portfolios are long (short) stocks of firms with the strongest (weakest) shareholder rights and yielded an 8.5% annual alpha during the 1990's. Subsequently, Core et al. (2006) tested the effect of the G-index on accounting returns and if this effect was not well recognized by the market, they found the G-index to be negatively related to ROA, stock returns, earnings announcement returns, and analyst forecast errors thus supporting their argument of mispriced information. These results are confirmed by Bebchuk et al (2009) who create an entrenchment index (E-index) that uses only the most relevant governance indicators.

However, if mispriced ESG information exists, it should eventually be corrected for when the market recognizes the benefits of for instance the CSR investments by firms. This will not happen overnight but is more likely a gradual learning process. Bebchuk et al. (2013) show that the market learned to correctly price governance information somewhere around 2001 after an increased supply of governance information in for instance academic and news articles. Although governance was at first significantly related to abnormal stock returns, earnings announcement returns, and analyst forecast errors, these relationships decreased to become statistically and economically insignificant after this flow of information. In Chapter 2, I test if environmental and social information is or has been mispriced by the market by performing mispricing tests as in Core et al. 2006. I also test for the possibility that the market has learned about ESG information.

In sum, there is empirical evidence that supports the positive effects on performance of mispriced ESG information, some of it direct although mostly indirect. Over different timespans and using different screening methods, researchers have shown that portfolios of stocks selected using ESG information *can* generate abnormally high risk-adjusted returns. However, even if ESG information is mispriced and an investor can benefit from this information, this benefit should eventually cease to exist as the market learns to correctly price this information. With this notion I make the bridge to the fifth Chapter of this dissertation. Reaping the benefits from SRI implies having a long investment horizon since learning does generally not happen overnight, and the investment universe is smaller resulting in less alternative investment opportunities. In Chapter 5 I test if institutions with tastes against controversial assets respond differently to permanent (cash flow) news as compared to temporary (discount rate) news components.

Summary of the papers in the dissertation.

In Chapter 2 I present a paper on the mispricing of stakeholder relations information in stock prices. This paper adds to the literature that tries to understand if trading based on stakeholder related information is *profitable*. In Chapters 3 and 4 I look at the effects of the *values* based SR investment style by analyzing the effects of excluding controversial stocks from investment portfolios of mutual funds (Chapter 3). And investigate the social and environmental investment preferences of Dutch pension fund participants as well as the consequences of offering customized investments to pension beneficiaries (Chapter 4). Finally, I analyze how preferences against assets with environmental and social controversies translate into trading behavior of institutional investors. Chapters 2 and 3 are joint work with my supervisors Jeroen Derwall, Kees Koedijk and Jenke Ter Horst. The fourth Chapter is with Rachel Pownall. And the fifth Chapter is with Jeroen Derwall, Nadja Guenster and Paulo Rodriguez.

2. Stakeholder relations and stock returns: on errors in investors' expectations and learning

In Chapter 2 we analyze stakeholder relations and stock returns because a significant number of institutional investors publicly state the belief that corporate stakeholder relations are associated with firm value in a manner that the financial market fails to understand. We investigate whether stakeholder information predicted risk-adjusted returns due to errors in investors' expectations. However, if this relationship exists, it should ultimately cease to do so when attention for such information increases. We build a stakeholder-relations index (*SI*) for a wide range of U.S. firms over the period 1992-2009 and provide evidence that the *SI* explained errors in investors' expectations about firms' future earnings. Then we apply a Quandt (1960) breakpoint analysis on stock returns of *SI* based investment portfolios to identify a point in time after which the market learned to correctly price stakeholder relations.

We find that this breakpoint is around the first quarter of 2004 and assure it is in line with the availability of information on stakeholder related business practices measured by shareholder proposals filed on stakeholder related issues and the number of firms issuing stand alone CSR reports.

We find that the *SI* was positively associated with long-term risk-adjusted returns over the period 1992-2004. Straightforward and naïve trading strategies generate risk adjusted abnormal returns of 3.5 to 5.5% annually. While similar equity portfolios lead to non significant abnormal returns over the period thereafter. Since abnormal performance might also stem from other factors than errors in investors' expectations we show in explicit mispricing tests that the *SI* was positively related to earnings announcement returns and errors in analysts' earnings forecasts over the period 1992-2004. However, as attention for stakeholder issues became more widespread, subsequently, these relationships diminished considerably. These results are consistent with the idea that increased investor attention for stakeholder issues eventually eliminates mispricing. Therefore, investors cannot justify using stakeholder related information as a source of sustainable outperformance.

3. Can investors profit from social tastes? Evidence from mutual fund holdings

When tastes affect investment decisions of a significant number of investors they have the potential to affect asset prices and consequently also expected returns (Fama and French 2007). In this paper we evaluate whether tastes for socially sensitive stocks affect holdings of US equity mutual funds. We start with a comparison of socially responsible investment funds to conventional funds and document on the existence of conventional funds that have "more socially responsible" holdings than SRI labeled funds. More specifically, we find that on average SRI funds behave "more" socially responsible, yet in a more in depth comparison we find that 10% to 30% of all US equity mutual funds is less (more) exposed to controversial

firms or firms with many ESG concerns (strengths).

Subsequently we analyze whether these exposures to socially sensitive stocks affect mutual fund performance. Our findings indicate that especially investments in Tobacco, Alcohol, and Gambling stocks have the potential to positively affect risk-adjusted fund returns, while exposures to the most socially responsible firms negatively affect performance. This potential is not fully exploited by the mutual funds in our sample as they hold diversified portfolios resulting in small exposure differences between funds. These small exposure differences also explain why the literature has generally found no performance differences between SRI labeled and conventional funds. Based on our main findings we advice the use of holdings based analyses when investigating the effects of social tastes on investment portfolios.

4. Attitudes towards socially and environmentally responsible investment

In the 4th Chapter we look at SR investing from a different angle than the previous two Chapters that investigate returns. The Chapter adds to the question on how financial institutions can serve their clients with SR investment products. We take a deeper look into pension investment products that have the potential to suit individuals' social norms. This is important since institutional investors invest billions of dollars on behalf of investors whilst knowing little about investors' social preferences.

Motivated by risk adjusted returns, legislative difficulties and evidence of financial illiteracy on a large scale, pension funds do not provide differentiated funds to meet the values based investment style which beneficiaries desire. Using data from a customized wave of the Dutch CentERdata panel for citizens who are required to participate in a pension plan, we find significant variation in stated preferences towards proposed social investment screens.

Subsequently we show that although individuals are able to express their preferences towards social investment criteria they are not able to translate these values into investment decisions consistently. This is partially driven by the low financial sophistication of households. Finally, to emphasize the importance of these findings we show that the majority of beneficiaries derive positive utility from environmental and social pension investment screens and that expressing a preference towards screened pension investments is the most important driver of this effect.

Important policy implications follow from this Chapter. First of all we show that pension fund participants on average care about how their money is invested. Furthermore, they can have contradictive preferences; therefore, some form of customization can be utility enhancing. However, giving beneficiaries any freedom of social and environmental choice that at the same time has a financial impact is not a feasible solution.

5. Values and investments: Evidence from institutional trading responses to news components

The 5th and final Chapter reports on an analysis of how institutions trade on different kinds of information about the firms they invest in. We know from the literature that norms and values driven investment styles lead to investment exclusion of firms that operate in controversial business areas or firms with environmental controversies (e.g. Heinkel et al. 2001, and Hong and Kacperczyk 2009). The literature has linked these exclusions to norms and values driven investment styles (see Hong and Kacperczyk 2009, and Hong and Kostovetsky 2012). In this Chapter we use the holdings of financial institutions to identify socially conscious investors, those with relatively low investments in assets with social controversies. We argue that social preferences do not only manifest itself as controversy avoidance in the stock selection process but also influence the investment decisions for all stocks in the portfolio on average. In line with the United Nations-backed Principles for

Responsible Investment and Nevins, Bearden and Money (2007) we argue that socially conscious investors are more long term oriented and are therefore less responsive to stock return news that is not related to firm fundamentals.

To test these predictions we use a vector autoregression (VAR) to separate returns into a *cash-flow news*- and an *expected return news* component following Vuoltheenaho (2002). We add institutional ownership data to the VAR as in Cohen et al. (2002) and separate institutional ownership into *socially conscious* institutional ownership and *conventional* institutional ownership. This enables us to observe if *socially conscious* institutions trade differently than do conventional institutions given news events occur.

Our findings indicate that *socially conscious* investors respond significantly less to expected return (non-persistent) news while these investors respond stronger to cash-flow (persistent) news. This evidence is in line with the view that social values do not only influence tastes against stocks with social controversies but also the investment decisions for all assets in the portfolio on average.

Chapter 2

2. Stakeholder relations and stock returns: on errors in investors' expectations and learning[8]

2.1. Introduction

"… we believe that, in the long run, an investment approach that identifies and invests in companies with sustainable business models serves shareholders best. Towards that end, we have developed a process that combines thorough financial analysis with another, critically important set of factors that most investment managers ignore…"

(PAX World Investments[9])

Financial institutions spend considerable time aligning their investment goals with the well-being of non-financial stakeholders and the community at large, by integrating environmental, social, and corporate governance (ESG) criteria with their investment decisions. Almost all institutions publicly justify those investments based on the argument that ESG information positively contributes to their investment performance. For example, more than 850 institutional investors worldwide, representing about $25 trillion assets under management, are signatories of the United Nations-backed Principles for Responsible Investing (PRI). According to PRI, institutional investors have a fiduciary duty to act in the long-term interests of beneficiaries, and ESG factors are relevant in this context because of their effect on the performance of investment portfolios.[10] Many of these investors are enamored with the

[8] This Chapter also circulates as Borgers, Derwall, Koedijk, Ter Horst (2013).

[9] http://www.paxworld.com/investment-approach/ (retrieved in 2010)

[10] See for example http://www.unpri.org/about-pri/the-six-principles/

idea that when firms improve their stakeholder relations they create intangible long-run economic benefits that are neither adequately reflected in firms' financial statements nor properly valued by the capital market.

This performance-oriented motivation for integrating stakeholder information into investments is nevertheless ambitious and remarkable. The notion that such information provides investors with a long-term competitive advantage goes against conventional economic wisdom and a large body of empirical evidence that active investors fail to beat the market consistently (e.g., Carhart 1997).[11] Even if better stakeholder relations are associated with higher future earnings in a manner that the market has not properly understood, economic logic predicts that such information provides investors with a competitive advantage *in the short-run, but not in the long-run.* Both theory and empirical evidence indicate that the documentation of profitable investment opportunities attracts investor attention and eventually contributes to market efficiency (e.g. Schwert 2003, Chordia, Subrahmanyam, and Tong 2012, Bebchuk, Cohen, and Wang 2013). Superior risk-adjusted returns that investors can earn by exploiting "mispriced" information, if any, should eventually cease to exist as the capital market learns and understands the earnings implications of this information.

This paper provides evidence that the quality of stakeholder relations originally did convey information about future risk-adjusted returns due to errors in investors' expectations, but less so as soon as the capital market paid more attention to stakeholder issues. The evidence on expectational errors is based on three common analyses that are considered complements in empirical studies on stock market anomalies (see Core, Guay, and Rusticus (2006); Edmans (2011); Bebchuk et al. (2013)). We first construct an annual stakeholder-

[11] Moreover, equilibrium models of asset prices predict that firms with strong stakeholder relations may even have lower expected returns if socially responsible investors drive up their stock prices (see, Heinkel, Kraus, and Zechner 2001, Hong and Kacperczyk 2009).

relations index (*SI*) for U.S. firms and then estimate risk-adjusted returns on stock portfolios that are formed using the *SI* over the period 1992-2009. We subsequently investigate whether stakeholder information predicts future earnings announcement returns. We complement these studies with an analysis of the association between stakeholder relations and errors in analysts' forecasts of firms' long-term earnings growth.

While our analyses suggest that stakeholder information was associated with risk-adjusted returns because of unexpected earnings, they also point out that evidence of errors in investors' expectations has weakened in recent times. While the *SI* positively related to risk-adjusted portfolio returns, earnings announcement returns, and analysts' long-term forecast errors over the period 1992-2004, these relationships diminished once stakeholder issues arguably attracted substantially greater attention in the capital market.

The conclusion that follows from the analyses is consistent with the learning hypothesis of Bebchuk et al. (2013), and has implications for those institutional investors that pursue both financial and social goals. On the one hand, the results imply that a performance-oriented investment case for integrating stakeholder issues in investment decisions has weaker empirical foundations than before, at least when it leans on easily obtainable information and rather elementary trading rules. But the conclusion that stakeholder management nowadays does not contribute to errors in expectations incentivizes company managers to place stakeholder issues higher on the corporate agenda. The results also expand on those studies on socially responsible investing (SRI) that present evidence to support the notion that certain stakeholder information is mispriced.[12] Especially Edmans (2011) presents comprehensive evidence that the stock market does not entirely value the intangible assets that companies create through strong relations with their employees. Our results suggest that

[12] See Derwall, Guenster, Bauer, and Koedijk (2005), Kempf and Osthoff (2007), Galema, Plantinga, and Scholtens (2008), Statman and Glushkov (2009), Edmans (2011), Derwall, Koedijk, and Ter Horst (2011).

such "mispricing" has diminished over time as the capital market eventually learned about the implications of stakeholder relations for corporate valuation.

This study proceeds as follows. The theoretical foundations of this study are discussed in Section 2.2 of the Chapter. Section 2.3 describes the data and variables that we use to measure the quality of stakeholder relations. Section 2.4 covers the main empirical analyses, and Section 2.5 discusses additional tests. Section 2.6 concludes this study.

2.2. Theoretical background

The idea that firms with better stakeholder relations have higher future earnings can be justified by both instrumental stakeholder theory (e.g. Cornell and Shapiro 1987, Zingales 2000) and the resource based-view of the firm (e.g. Wernerfelt 1984, Barney 1991, Hart 1995, Russo and Fouts 1997). That these advantages are often intangible, not readily quantifiable, and materialize in the long-term provides investors in search of underpriced assets with one argument for integrating stakeholder information into investment decisions. Several institutional investors, such as various signatories of PRI, contend that financial markets do not appreciate these intangibles. For example, the Enhanced Analytics Initiative (EAI) is an investor initiative (now merged with PRI) that incentivizes analysts to routinely consider so-called "extra-financial information", so that their investment recommendations are improved (O'Loughlin and Thamotheram 2006). According to EAI, extra-financial factors are "*those which are likely to have at least a long-term effect on business results but which seldom get integrated into investment decisions*…", ranging from "*corporate governance and executive remuneration, to occupational health and safety and human capital practices, and to the environmental and social impacts of corporate activity*" (O'Loughlin and Thamotheram 2006, p. 6).

Whether such factors reflect intangibles that are not properly reflected in stock prices has also attracted considerable attention in empirical studies over the last years. On the academic front, several relatively recent studies have suggested that stocks of companies with better stakeholder relations have produced anomalously positive average returns in the U.S. stock market. See, for example, Derwall et al. 2005; Kempf and Osthoff 2007; Statman and Glushkov 2009, Edmans 2011; Derwall et al. 2011. In particular Edmans (2011) showed that companies with stronger employee satisfaction not only had higher risk-adjusted returns in the stock market but also exhibited both higher earnings announcement returns and higher long-term earnings surprises.

As Derwall, Koedijk and Ter Horst (2011) point out, if these findings indeed reflect mispricing, then it is questionable that they will persist in the long run. Standard economic theory predicts that mispriced information eventually disappears as investors learn about the anomaly. Prior studies provide evidence that many widely publicized anomaly variables were able to predict stock returns during the sample period in which they were first identified, but less so after their discovery (e.g. Schwert 2003, Chordia, Subrahmanyam, and Tong 2012). There are at least two reasons to expect that the capital market has come to better understand the value-relevance of corporate stakeholder relations.

First, anecdotal evidence points out that investor attention for stakeholder issues has risen substantially in recent years. Industry surveys consistently conclude that the amount of assets managed by institutional investors that integrate so-called environmental, social and governance (ESG) issues has grown considerably over the last decade. For example, according to the U.S. social investment forum (2010), about 55 mutual funds (representing US$ 12 billion under management) integrated ESG factors into investment choices in 1995, while almost 500 funds with US$ 569 billion under management employed such investment criteria in 2010. Outside the U.S., several investor initiatives, such as EAI in 2004 and PRI in

2006 contributed to the worldwide mainstreaming of ESG, encouraging mainstream investors to routinely integrate stakeholder issues with investment decisions.[13]

Second, in a closely related study, Bebchuk et al. (2013) show that the corporate governance index of Gompers, Ishii and Metrick (2003) originally related significantly to risk-adjusted stock returns, analysts' earnings forecast errors, and abnormal earnings announcement returns—but not after 2001, when governance issues attracted structurally greater attention among financial media, academic studies, and shareholder proposals issued by institutional investors. Consequently, they conclude that investors learned about the association between governance indexes and firms' profitability as a result of this heightened attention for corporate governance. The conclusion of Bebchuk et al. (2013) has potentially important implications for our study because many investors learn about the value-relevance of governance issues in tandem with stakeholder issues, in particular via the ESG acronym.

In summary, the growth of investors who employ corporate stakeholder information for pursuing the goal of superior returns raises two empirical questions. The first question addressed in this paper is whether there is justification for the belief that errors in expectations causes firms' stock returns to be associated with the quality of stakeholder relations ("the errors-in-expectations hypothesis"). If so, the natural follow-up question is whether risk-adjusted stock returns stemming from errors in investors' expectations eventually cease to exist following investors' heightened attention for stakeholder information, in the spirit of the "learning hypothesis" of Bebchuk et al. (2013). The goal of

[13] The "ESG" acronym became widespread due to summits involving large investment companies, and is an explicit outcome of investors seeking to "mainstream" the use of stakeholder information by the investment community. For a review of alternative terminologies, see also Bessire and Onnée (2010).

this study is to investigate whether both hypotheses find support in analyses of risk-adjusted portfolio returns, earnings announcement returns, and errors in analysts' earnings forecasts.[14]

2.3. Evaluating corporate stakeholder relations

We annually evaluate firms' stakeholder relations using the STATS database from Kinder, Lydenberg and Domini and co. (KLD), which is the longest-running source of stakeholder information and used extensively by investors. STATS summarizes this information for mostly Standard & Poor's (S&P) 500 constituents as of 1991, the 1,000 largest publicly traded U.S. companies from 2001 to 2002, and the 3,000 largest publicly traded U.S. companies every year thereafter.

KLD specializes in evaluating firms regarding issues such as environmental performance (e.g. hazardous waste, regulatory problems, emissions and pollution prevention, and environmental management systems), community involvement (e.g. charitable and innovative giving, support for housing and education, and volunteer programs), diversity (e.g. women on the board of directors, CEO gender, the promotion or contracting of women and minorities, and work/life benefits), employee relations (e.g. workplace health and safety issues, workforce reductions, retirement benefits, worker involvement programs, and union relations), product quality (e.g. marketing-contracting concerns, product safety, and benefits to the economically disadvantaged), and human rights issues.[15] For each category, KLD

[14] In principle, heightened attention may also affect the demand for specific stocks, which may influence their returns. Edmans (2011) investigates whether increased demand for stocks of America's Best Companies to Work For explains these stocks' positive risk-adjusted returns, for which he finds little evidence.

[15] We adjusted the diversity measure to correct for KLD's overweighting of issues related to female representation by setting a maximum of 1 to the sum of all diversity issues related to female representation.

subjects every firm to a number of "strengths" and "concerns" indicators, with "1" ("0") indicating the presence (absence) of a strength or concern.[16]

For every firm we develop an aggregate stakeholder-relations index (henceforth, *SI*) on a yearly basis, using the strengths and concerns indicators from KLD. To construct the *SI*, we follow the common practice of adding all strengths and subtracting all concerns in a given year (see, e.g, Hong and Kostovetsky (2010) and Jiao (2010)). We omit from this procedure the indicators of human rights issues, because KLD did not cover these issues consistently throughout the sample period. Moreover, we industry adjust these scores by subtracting the mean score within an industry from the firms' score.[17]

From a statistical standpoint, the aggregate of the individual indicators has the most desirable distributional characteristics compared to disaggregate measures. For example, around 80 percent of all firm-year observations do not experience a single strength or concern in the areas of community involvement or environment, whereas this occurs only in 14 percent of the cases when all stakeholder categories are aggregated. Therefore, undesirable distributional features makes the use of too disaggregate measures problematic in common tests of errors in expectations.

Panel A in Table 2.1 presents summary statistics for the SI. The *SI* has a mean of zero and has a standard deviation of 1.68. Panel B reports correlations between the *SI* and a number of elementary financial variables based on data from Compustat, which creates a basic impression of the financial characteristics of firms with stronger stakeholder relations relative to those with weaker relations. These basic statistics support the popular belief that firms with better stakeholder relations tend to have larger accounting profits (e.g. Russo and

[16] Next to covering these strengths and concerns indicators, KLD offers a separate analysis of firms' involvement in controversial sectors, specifically, alcohol, gambling, firearms, military, nuclear power, and tobacco.

[17] We use the Fama French 10 industry definition to have sufficient within industry variation.

Fouts 1997, King and Lenox 2002, Jiao 2010), higher price-to-book ratios (Galema et al. 2008), and lower leverage ratios (Verwijmeren and Derwall 2010, Bae, Kang, and Wang 2011). Whether the *SI* is also associated with higher risk-adjusted returns in the stock market is central to the next Section of the paper.

Table 2.1 Summary statistics and correlations

Reported are descriptive statistics on the *SI* (Panel A), and pairwise correlations between the *SI* and several firm characteristics (Panel B). Reported in parentheses are the involved numbers of observations. Return on assets (ROA) is defined as the ratio of operating income (after depreciation and amortization) divided by total assets, the book-to-market equity defined as the book value of equity plus book value of deferred taxes divided by the market value of equity (common shares outstanding * share price at the end of the fiscal period). Leverage as long-term debt to total assets.

Panel A: Distributional characteristics of the SI

Variable	Obs.	Mean	St.Dev.	Min	Max
SI	22792	0.00	1.68	-9.28	9.18

Panel B: Correlation between the SI and firm characteristics

	ROA	Log Book/market	Log assets	Leverage	1-yr Sales growth
SI	0.038	-0.101	-0.079	-0.082	0.017
	(22792)	(22133)	(22512)	(22792)	(22390)

2.4. Empirical analysis

We present three complementary analyses of errors in investors' expectations that are common in studies on stock market anomalies. The first analysis revolves around risk-adjusted returns on investment portfolios that are formed based on the *SI*. The second analysis focuses on stock returns around firms' earnings announcements. The third analysis explores investors' expectations by means of analyst forecasts.

2.4.1. Portfolios and decreasing risk-adjusted returns

Our empirical analysis starts with an evaluation of the returns on stock portfolios that are formed using the *SI*. Our primary objectives in this Section of the study are to investigate (i) whether portfolios composed of stocks that ranked high on the *SI* earned a significantly higher risk-adjusted return than those that score lower on the *SI*, and if so, (ii) whether the difference in risk-adjusted return eventually diminished once investors paid more attention to stakeholder information.

Every year, starting in April 1992, we rank all available stocks on the *SI*, and then allocate those stocks that rank above a specific upper threshold level to a top-ranked portfolio and those that rank below a bottom threshold level to the bottom-ranked portfolio.[18] We exclude from the portfolio construction those stocks that belong to KLD's list of controversial businesses, because prior research explicitly attributes their returns to risk premiums instead of errors in expectations (see Hong and Kacperczyk (2009)). Using the CRSP returns database, we compute the monthly returns on the portfolios during the twelve consecutive months after formation until the portfolios are updated based on the latest *SI* values, and we subsequently evaluate the time-series of portfolio returns over the period April 1992-December 2009.

Following previous studies that document significant risk-adjusted returns associated with the quality of corporate stakeholder relations, we derive risk-adjusted returns from the Carhart (1997) four-factor regressions:

$$R_{i,t} - R_f = \alpha_i + \beta_{0,i}(R_{m,t} - R_{f,t+1}) + \beta_{1,i} SMB_t + \beta_{2,i} HML_t + \beta_{3,i} UMD_t + \varepsilon_{i,t} \quad (2.1)$$

[18] The starting year in the KLD STATS database is 1991, but KLD usually releases its statistics in the first quarter of the subsequent year.

where $R_{i,t}$ is the return on a portfolio that is formed based on the *SI*, $R_{m,t} - R_{f,t}$ is the return on a portfolio composed of all stocks from the NYSE/AMEX/Nasdaq exchanges minus the one-month T-Bill rate from Ibbotson Associates, SMB_t is the return difference between a small cap portfolio and a large cap portfolio, HML_t is the return difference between a "value" portfolio (with a high book/market value ratio) and a growth (low book/market value) portfolio, UMD_t is the return difference between a portfolio of the past 12-month return winners and a portfolio of the past 12-month losers. A large amount of literature consistently points out that the four factors, which are taken from the Kenneth French Data Library, are important in explaining the returns on equity portfolios that are formed using stakeholder information.[19]

Table 2.2 shows average risk-adjusted returns and four-factor factor loadings measured over the entire sample period (April 1992-December 2009) for a number of portfolios that are formed using the *SI*. The regression parameters are largely consistent with earlier studies that have documented risk-adjusted returns associated with several of KLD's indicators. A portfolio composed of either the top third, or top fourth, or top fifth of all stocks ranked by the *SI* earned a higher average annualized risk-adjusted return than its bottom-ranked counterpart. The performance difference is economically significant, and in two of the three reported cases statistically significant at the conventional levels of significance. Table 2.2 also shows that much of the performance difference between the top-ranked and bottom-ranked portfolios is largely due to positive risk-adjusted returns of top-ranked portfolios. The risk-adjusted return on bottom-ranked portfolios are not significantly different from zero.

[19] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. See Fama and French (1993) and Carhart (1997) for details on the construction of the four factors.

We now turn to time variation in the risk-adjusted return on portfolios formed using the *SI*. A visual inspection of rolling-window regressions involving specification (2.1) provides the first indication that risk-adjusted returns on portfolios constructed using the *SI* have weakened over time. Figure 2.1 shows that the equal-weighted risk-adjusted return on a portfolio that is long in the top one-third of stocks and short in the bottom-third was persistently positive for a substantial number of years but eventually decreased considerably.

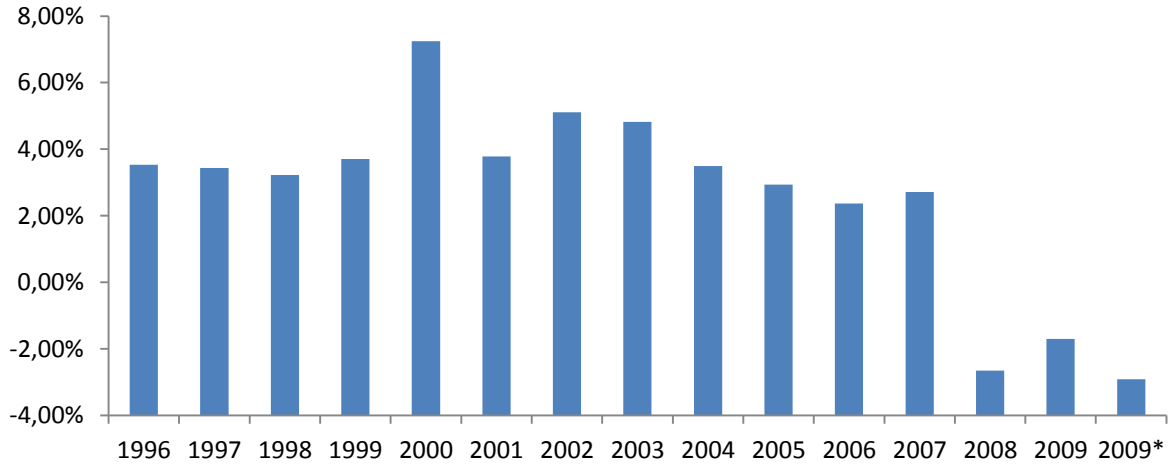
Table 2.2 Risk-Adjusted returns over 1992-2009

Every year, starting in April 1992, we rank stocks based on the *SI* and assign the top (bottom) third, fourth, or fifth of all ranked stocks to a top-ranked (bottom-ranked) portfolio. We run Carhart (1997) four-factor regressions to estimate risk-adjusted portfolio returns over the period April 1992-December 2009. Reported are annualized risk-adjusted returns and factor exposures for equal-weighted portfolios.

	Alpha	Rm-Rf	SMB	HML	UMD	R2
Top-minus-bottom third						
Top	2.46%**	1.03***	0.21***	0.44***	-0.17***	0.94
	(2.27)	(41.84)	(4.89)	(11.70)	(-8.24)	
Bottom	0.89%	1.07***	0.24	0.45***	-0.23***	0.91
	(0.60)	(33.43)	(5.09)	(6.56)	(-7.46)	
Top-minus-bottom	1.57%	-0.04	-0.03	-0.01	0.06**	0.08
	(1.19)	(-1.38)	(-0.98)	(-0.16)	(2.05)	
Top-minus-bottom fourth						
Top	3.22%**	1.03***	0.16***	0.46***	-0.18***	0.91
	(2.36)	(33.10)	(2.89)	(10.16)	(-5.61)	
Bottom	-0.30%	1.07***	0.21***	0.51***	-0.21***	0.91
	(-0.21)	(36.28)	(4.45)	(8.61)	(-6.99)	
Top-minus-bottom	3.52%***	-0.04	-0.06*	-0.05	0.04	0.06
	(2.75)	(-1.48)	(-1.71)	(-1.01)	(1.16)	
Top-minus-bottom fifth						
Top	2.99%**	1.02***	0.16***	0.49***	-0.16***	0.91
	(2.13)	(32.87)	(2.98)	(11.22)	(-4.81)	
Bottom	0.10%	1.04***	0.22***	0.49***	-0.23***	0.90
	(0.06)	(31.82)	(4.74)	(7.23)	(-6.87)	
Top-minus-bottom	2.89%*	-0.01	-0.06	0.00	0.07	0.06
	(1.96)	(-0.42)	(-1.40)	(0.05)	(1.61)	

Figure 2.1 year-by-year difference in risk-adjusted return between top- and bottom-ranked portfolios

Every year, we perform Carhart (1997) four-factor regressions using monthly return differences over the last 4-years between the portfolio composed of the top third of stocks ranked on the stakeholder relations index and the bottom-ranked counterpart. Reported are the annualized yearly risk-adjusted returns derived from equal-weighted portfolios ending in March of the year (except for 2009* which ends in December). The stakeholder-relations index *SI* is based on the sum of all strengths a firm receives in the areas of environment, community, diversity, employee relations, and product quality minus to sum of all concerns.



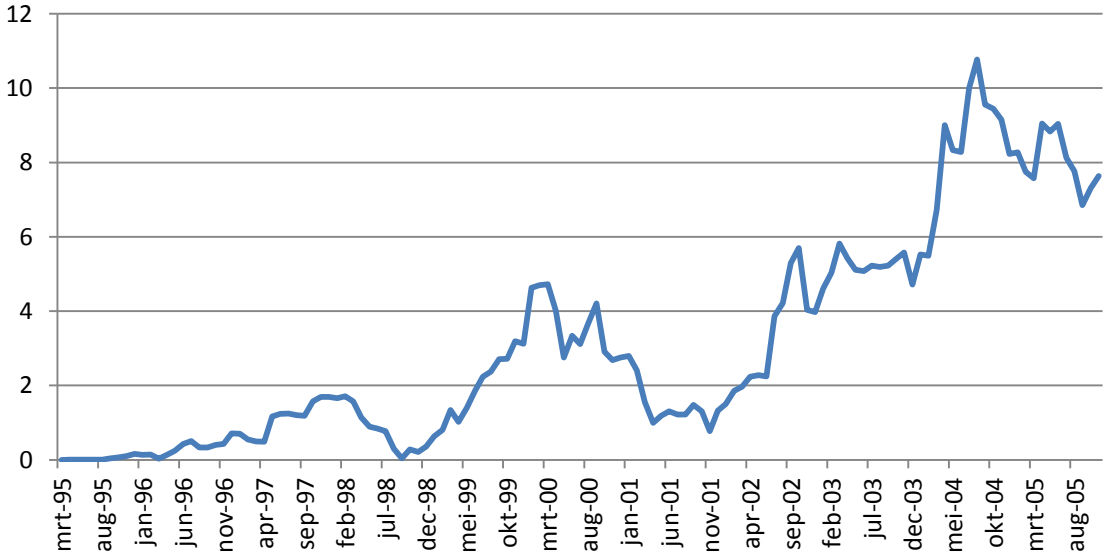
To explore more formally the time-variation in returns, we adopt a variant of the procedure described in Quandt (1960) and Bebchuk et al. (2013). The goal of the procedure is to identify a date that marks a structural break in risk-adjusted returns of portfolios that are formed based on the *SI*. The date that is identified in this way marks a break in the sense that the risk-adjusted returns across the two periods differ the most from a statistical point of view. To determine the break date, we estimate a variant of the Carhart (1997) regression, which allows risk-adjusted returns and portfolio factor loadings to vary across two periods.

$$R_{top,t} - R_{bottom,t} = \alpha_i * Post_t + \beta_{1,i}(R_{m,t} - R_{f,t}) * Post_t + \beta_{2,i} SMB_t * Post_t + \ldots \\ \ldots \beta_{3,i} HML_t * Post_t + \beta_{4,i} UMD_t * Post_t + \varepsilon_{i,t} \quad (2.2)$$

where $R_{top,t}$ is the return in month t on the top-ranked portfolio, $R_{bottom,t}$ is the return on the bottom-ranked portfolio, and the dummy variable $Post_t$ is an indicator variable that captures all months including and after a breakpoint date. To determine which break date marks the largest difference in risk-adjusted return between two periods, we re-estimate the model based on all possible variations of the indicator variable $Post_t$. Like Bebchuk et al. (2013), we compute the F-statistic on the coefficient on $\alpha * Post_t$ for each regression and then determine the break date from the regression that yields the largest F-statistic for this coefficient.

Figure 2.2 F-statistics from Quandt test on portfolio returns

Every year, starting in April 1992, we rank stocks based on the stakeholder-relations index (*SI*) and assign the top (bottom) third of ranked stocks to an equal-weighted top-ranked (bottom-ranked) portfolio. We apply a Quandt (1960) procedure to determine the date of a break in the risk-adjusted return difference between the two portfolios, which requires estimations model (2.2) using monthly returns from April 1992 to December 2009. Our Quandt test involves a re-estimation of model (2.2) based on all possible variations of the indicator variable *Post*. We impose that *Post* cannot equal 1 for the first 36 and last 36 months of our sample period in order to ensure that all factor loadings can be estimated properly. We compute the F-statistic on the coefficient on α*Post* for each regression, and identify the break date from the regression that yields the largest F-statistic for this coefficient.



In Figure 2.2, we give a graphical example of one specific Quandt test result that is relevant for break date determination. The F-statistics suggest that August 2004 marks a break in the return difference between the equal-weight top-ranked and bottom-ranked portfolio. For this month, the F-statistic on $\alpha^* Post_t$ is 10.77, which is almost twice as large as the F-statistic corresponding to the same month one year earlier, and about nine times as large as the F-statistic observed 3 years earlier.

We apply this procedure to determine break dates for a number of "top-minus-bottom" ranked portfolios that can be formed using the *SI*, and then measure risk-adjusted return before and after the break-date. In independent analyses, we allow the top and bottom portfolios to comprise either the top (bottom) third, fourth, or fifth of all stocks that are ranked on the *SI*. Table 2.3 shows the risk-adjusted returns on both equal-weighted and value-weighted portfolios, measured over, respectively, the full sample period, the pre-break period, and the post-break period.

Concerning equal-weighted returns, the Quandt test marks as break dates, respectively August 2004 for top-minus-bottom third portfolios, July 2004 for top-minus-bottom fourth portfolios, and October 2002 for top-minus-bottom-fifth portfolios. As for value-weighted returns, the corresponding dates are November 2005, February 2003, and November 2005. The average date, then, corresponds to June 2004.

Finding breakpoints close to 2004 seems consistent with indicators of attention to stakeholder issues among companies and investors. For example, in order to explore a proxy for attention by investors, we counted the yearly number of shareholder proposals on corporate social policy issues that were mainly (co)sponsored by institutions from 1991 onwards (after removing proposals from individuals, religious groups, special interest groups, and unknown sponsors). We derived these results from an analysis of the RiskMetrics database of shareholder proposals in the U.S., which involves mostly S&P 1500 constituents.

What becomes apparent in Figure 2.3 is that firms received structurally more proposals on social policy issues in recent years. Also pointing to heightened attention for stakeholder issues is the increasing volume of information that U.S. companies disclose on stakeholder relations. Dahliwal, Li, Tsang, and Yang (2011) investigated the number of U.S. firms per year that voluntarily disclosed CSR information. Their results suggest that aggregate CSR reporting increased substantially, first temporarily in 2001 and then more permanently from 2003 onwards.

Figure 2.3: Number of shareholder proposals on stakeholder issues

We collect all shareholder proposals involving S&P 1500 firms from Riskmetrics over the period 1991-2008. For each proposal, we identify the (co)sponsor and eliminate proposals that are exclusively sponsored by individuals, religious institutions, and special interest groups (e.g. PETA, Friends of the Earth). To identify stakeholder issues we take all shareholder proposals that Riskmetrics classifies as social policy issues ("SRI") and add all "crossover" proposals, i.e., proposals involving social issues that investors submit tied to executive compensation.

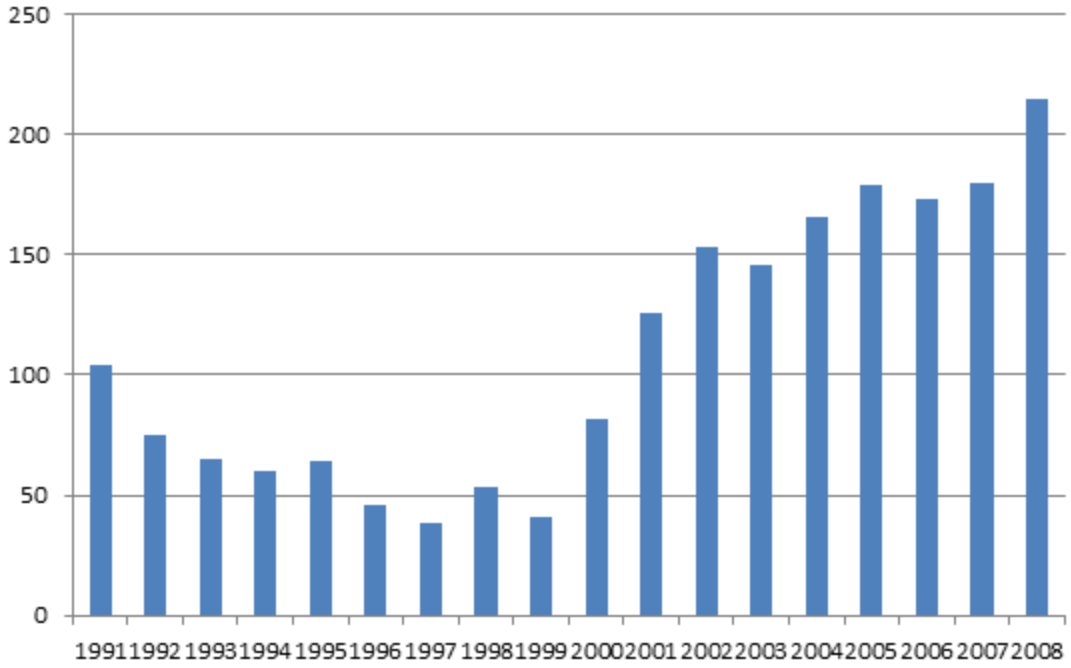


Table 2.3 Quandt test on difference in risk-adjusted returns over time for portfolios based on the *SI*

Every year, starting in April 1992, we rank stocks based on the stakeholder-relations index (*SI*) and assign top-ranked (bottom-ranked) stocks to an equal-weighted or value-weighted top-ranked (bottom-ranked) portfolio. We explore alternative stock selection rules: top and bottom third, fourth, or fifth of all stocks ranked on the *SI*. We apply a Quandt (1960) procedure to determine the date of a break in the risk-adjusted return difference between the portfolios. We use monthly returns from April 1992 to December 2009. We re-estimate model (2.2) based on all possible variations of the indicator variable *Post*. We compute the F-statistic on the coefficient on α**Post* for each regression, and identify the break date from the regression that yields the largest F-statistic for this coefficient. We impose that *Post* cannot equal 1 for the first 36 months and last 36 months of our sample period in order to ensure that all factor loadings can be estimated properly. Based on the break dates, we estimate model (2.1) for the returns on Top-, Bottom, and Top minus bottom ranked portfolios, before the breakpoint date and after breakpoint date. The first column reports risk-adjusted returns measured over the entire sample period April 1992-December 2009.*, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

	Equal-weighted α				Value-weighted α			
SI portfolio	1992 - 2009	Break date	Pre-break	Post-break	1992-2009	Break date	Pre-break	Post-break
Top minus bottom third	1.57%	Aug-04	4.19%***	-2.76%*	2.02%	Nov-05	3.71%**	-2.60%
	(1.19)		(2.81)	(-1.82)	(1.26)		(2.06)	(-1.11)
Top	2.46%**		4.26%***	-1.31%	1.21%		2.33%*	-1.85%
	(2.27)		(3.53)	(-1.02)	(1.10)		(1.80)	(-1.19)
Bottom	0.89%		0.08%	1.45%	-0.81%		-1.38%	0.76%
	(0.60)		(0.04)	(1.35)	(-0.73)		(-1.03)	(0.59)
Top minus bottom fourth	3.52%***	Jul-04	5.52%***	0.33%	2.96%*	Feb-03	5.48%**	-0.74%
	(2.75)		(3.26)	(0.24)	(1.80)		-2.41	(-0.34)
Top	3.22%**		5.00%***	0.35%	1.80%		5.54%***	-1.45%
	(2.36)		(3.17)	(0.26)	(1.36)		-2.79	(-0.99)
Bottom	-0.30%		-0.52%	0.02%	-1.17%		0.06%	-0.71%
	(-0.21)		(-0.30)	(0.02)	(-1.02)		(0.034)	(-0.55)
Top minus bottom fifth	2.89%*	Oct-02	5.04%***	0.13%	3.01%*	Nov-05	5.45%***	-2.16%
	(1.96)		(2.78)	(0.07)	(1.71)		(2.68)	(-0.83)
Top	2.99%**		4.27%**	-0.34%	1.67%		4.40%***	-1.60%
	(2.13)		(2.45)	(-0.25)	(1.26)		(2.74)	(-0.81)
Bottom	0.10%		-0.77%	-0.47%	-1.34%		-1.04%	0.57%
	(0.06)		(-0.40)	(-0.34)	(-1.11)		(-0.62)	(0.40)

Using the dates determined by this test, we see in Table 2.3 that the average risk-adjusted return differences between top- and bottom-ranked portfolios are positive, economically large, and statistically significant prior to each break date. In contrast, the post-break risk-adjusted return is not significantly different from zero in five out of the six analyses, and negative (albeit significant at the 10% level) in one case.

Based on the average of the different Quandt test results, it stands to reason that the quality of stakeholder relations at first related positively to (risk-adjusted) stock returns, but that such a relation has decreased or diminished as from 2004. Because KLD tends to report its yearly evaluation of firms' stakeholder relations in the first quarter of the next year, we would expect that KLD's indicators released after the first quarter of 2004 conveys less information about risk-adjusted returns than indicators released in the years before. For this reason, we report in Table 2.4 the difference in risk-adjusted return between top-ranked portfolios and their bottom-ranked counterparts during, respectively, the periods April 1994-March 2004 and April 2004-December 2009.[20]

The results in Table 2.4 further corroborate that those trading rules based on the *SI* that produced a positive risk-adjusted return have done so significantly only in the first sub-period. All equal-weighted and value-weighted portfolios that score high on *SI* significantly outperformed their bottom-ranked counterparts during the period April 1992-March 2004, but most of these portfolios ceased to exhibit significant differential risk-adjusted returns during the period April 2004- December 2009.

At first glance, the results presented in this Section suggest that the financial market has temporarily been too pessimistic about the value-relevance of stakeholder performance, leading to positive risk-adjusted returns, but then learned about the earnings difference among firms that differ in the quality of stakeholder relations. However, because long-term

[20] We also performed all analyses using 2003 as the breakpoint year. These results are available upon request.

risk-adjusted returns can also emerge for reasons other than "mispricing", we now turn to more explicit tests of errors in investors' expectations. [21]

Table 2.4 Difference in risk-adjusted return over time: Before and after April 2004.

Every year, starting in April 1992, we rank stocks based on the stakeholder-relations index (*SI*). We then assign stocks to either an equal-weighted or a value-weighted top-ranked (bottom-ranked) portfolio. We run Carhart (1997) four-factor regressions as in model (2.1) to estimate the difference in risk-adjusted return between the portfolios over two consecutive periods April 1992-March 2004 and April 2004-December 2009. We explore alternative stock selection rules: top minus bottom third, fourth, and fifth of stocks ranked on the *SI*. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

	Equal-weighted α			Value-weighted α		
SI portfolio	1992-2009	92-04	04-09	1992-2009	92-04	04-09
Top minus bottom third	1.57%	3.52%**	-2.30%	2.02%	3.43%*	-1.28%
	(1.19)	(2.44)	(-1.58)	(1.26)	(1.72)	(-0.63)
Top minus bottom fourth	3.52%***	5.24%***	0.85%	2.96%*	4.36%**	-0.30%
	(2.75)	(3.08)	(0.60)	(1.80)	(2.05)	(-0.14)
Top minus bottom fifth	2.89%*	4.36%**	1.33%	3.01%*	4.42%**	-0.24%
	(1.96)	(2.59)	(0.74)	(1.71)	(2.00)	(-0.11)

2.4.2. Earnings announcement returns

Researchers have suggested that stock returns around earnings announcements can be used to detect more explicitly errors in investors' expectations investors' concerning firms' earnings.[22] In this section, we study abnormal earnings announcement returns to determine the extent to which the time-variation in average risk-adjusted returns on the aforementioned *SI*-based strategies represent investors' initial misunderstanding and subsequent learning

[21] For example, risk-adjusted stock returns may alternatively stem from risk premiums that are overlooked by models that researchers use to determine expected returns (see, e.g., Fama and French (1993)), and from data snooping (Lo and MacKinlay (1990)).

[22] See for example Chan, Jegadeesh and Lakonishok (1996), Sloan (1996), La Porta, Lakonishok, Shleifer, and Vishny (1997), Core et al. (2006), and Bebchuk et al. (2013).

about firms' earnings. If it is true that firms with higher *SI* values realized higher profits than anticipated by investors, we would expect that investors' surprises are reflected in higher abnormal returns around earnings announcements. We would also expect that the *SI* ceases to explain earnings announcement returns in times of heightened capital market attention for stakeholder issues.

We perform an event study to measure firms' stock returns around the announcements, using quarterly earnings announcement dates from I/B/E/S and daily stock returns from CRSP. For each stock, we compute daily abnormal returns from various days before until one day after each announcement, where the daily abnormal return (AR) is the difference between the realized return and the return predicted by the Carhart (1997) four-factor model. The return prediction model is re-estimated for each firm before every earnings announcement, using stock returns observed over a 250-day period that ends 20 days before the announcement date. The daily abnormal returns are subsequently converted to cumulative abnormal returns (CARs) over, respectively, three-day (-1,1), five-day (-3,1), seven-day (-5,1), and twelve-day (-10,1) windows.

In the tradition of Bebchuk et al. (2013), we derive time-variation in the relation between the earnings announcement CAR and corporate stakeholder relations from pooled regressions that take the form:

$$\begin{aligned} CAR_{i,(tq-s,tq+1)} = {} & \alpha + \beta_1 SI_{i,t-1} + \beta_2 SI_{i,t-1} Subsample2_t + \beta_3 Subsample2_t + \ldots \\ & \ldots \sum_{k=1}^{K} \gamma_k Controls_{i,k,t-1} + \varepsilon_{i,tq} \quad for\ s \in \{1,3,5,10\} \end{aligned} \tag{2.3}$$

where $CAR_{i,(tq-s,tq+1)}$ is the cumulative (s+2)-day abnormal return around the earnings announcement for firm *i* in quarter *q* of year *t*. The vector of controls includes a dummy variable that captures firms' presence on KLD's list of controversial businesses and industry dummy variables. Of primary interest to us is the stakeholder-relations index *SI* and its

interaction with a dummy variable *Subsample 2* that equals 1 if earnings announcements occurred after March 2004, the period after which we expect that information from KLD conveys less information about errors in expectations than before (also see Table 2.4).

The estimated relationships between the *SI* and the earnings announcement CARs are reported in Table 2.5. All coefficients are multiplied by 1000 for expositional convenience. The regression results are consistent with the idea that better stakeholder relations was associated with higher risk-adjusted stock returns over the period 1992-2004 due to errors in investors' expectations. The coefficients concerning the *SI* point to a statistically and economically significant relationship with cumulative earnings announcement returns, regardless of the event window that we consider. For example, a one-point increase in *SI* is associated with roughly a 0.09 percent five-day abnormal return per quarterly earnings announcement during the period 1992-2004, which is equivalent to an annualized abnormal announcement return of about 0.36 percent. The average difference in *SI* score between the top one-third and bottom one-third bottom-ranked portfolio over this period is 4.33 (not tabulated), which multiplied with the estimated earnings announcement effect, implies an industry-adjusted difference in abnormal earnings announcement return of 1.56 percent.

Table 2.5 suggests not only that earnings announcement effects explain risk-adjusted returns associated with the *SI* over the period 1992-2004 but also that such earnings announcement effects have decreased subsequently. Independent of the event window, the coefficient on *SI*Subsample 2* is consistently negative and significant below the 5% significance level of significance. According to F-tests regarding the sum of the coefficients on *SI* and *SI*Subsample 2*, the decrease in the earnings announcement effect measured over 2004-2009 is large enough to make the positive earnings announcement effect in the earlier period disappear. None of the F-statistics rejects the null of a zero relation between the *SI* and earnings announcement CARs during the period 2004-2009. The decreasing relation between

the *SI* and these CARs over time is consistent with the notion that risk-adjusted returns associated with stakeholder information eventually disappear as rising attention causes investors to learn about the differential future earnings among firms with different stakeholder relations.

Table 2.5 Stakeholder relations and earnings announcement returns

We estimate the relationship between the stakeholder-relations index and cumulative earnings announcement returns using as in model (2.3). The dependent variable is the cumulative abnormal return (CAR) realized during periods of varying lengths around the quarterly earnings announcement date of each firm. We explain the CAR by the *SI* which is the stakeholder-relations index, Subsample 2 which is a dummy variable that equals 1 when earnings announcements occurred during the period April 2004-December 2009 and zero otherwise controls that includes a dummy variable that captures firms' presence on KLD's list of controversial businesses, and industry fixed effects based on the 48 industry classifications from the Kenneth French Data Library. In four independent regressions, we analyze the effect of stakeholder relations on CAR measured over, respectively three-day (-1,1), five-day (-3,1), seven-day (-5,1), and twelve-day (-10,1) event windows. The *t*-statistics (in parentheses) are derived from two-way clustered standard errors. The reported coefficients are multiplied by 1000 for expositional convenience. The F-test measures for each regression whether the sum of the coefficients on *SI* and *SI*Subsample 2* are different from zero. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

	Event window (days before, after)			
	-1,+1	-3,+1	-5,+1	-10,+1
SI	0.889***	0.895***	0.758**	0.803**
	(3.15)	(3.20)	(2.38)	(2.12)
*SI**Subsample 2	-1.032**	-1.140***	-1.000**	-1.338**
	(-2.46)	(-2.58)	(-1.98)	(-2.15)
Subsample 2	-0.722	-0.477	0.318	1.311
	(-0.60)	(-0.29)	(0.16)	(0.48)
Controversial business	2.320**	2.124*	1.147	1.246
	(2.08)	(1.67)	(0.76)	(0.63)
Constant	10.829***	6.575	3.714	1.712
	(3.05)	(1.45)	(0.76)	(0.36)
Observations	78,340	78,323	78,319	78,310
Adj. R-squared	0.002	0.001	0.001	0.002
F-test ($\beta1+\beta2=0$)	0.220	0.560	0.411	1.356
Prob. > F	0.639	0.454	0.521	0.244

Given that stakeholder issues attracted substantial attention in recent years as part of a broader interest in environmental, social and corporate governance issues, one might ask whether the diminishing relation between the *SI* and earnings announcement CAR is driven by the learning effect that Bebchuk et al. document for certain corporate governance issues. To ensure that the learning effect documented in our study is unique, we also run regressions after expanding the vector of control variables with the corporate governance indexes that Bebchuk et al. associate with their learning hypothesis. The first governance index measures the number of anti-takeover provisions (*G Index*) developed by Gompers, Ishii, and Metrick (2003), with higher values for the index implying more provisions and hence a weaker corporate governance structure. The second index is the entrenchment index (*E Index*) of Bebchuk, Cohen and Ferell (2009), which is a subset of the *G Index* that has been shown to better predict firm value and abnormal stock returns. Both indexes were taken from the authors' websites.

Table 2.6 shows that the diminishing association between the *SI* and earnings announcement returns is present across all models that also contain the corporate governance indexes. The diminishing association between the *SI* and earnings announcement CAR not only continues to be significant in all of the models but also appears to be more robust than the time-variation in the relation between corporate governance and earnings announcement returns during our sample period. Specifically, the coefficient estimates for the *SI* and the *SI**Subsample 2 variables are highly stable and statistically significant regardless of the model employed, whereas only the coefficient on the entrenchment index differs significantly from zero in the first subsample period, at the 10% level.

Combined with the results from the previous section, the earnings announcement regressions yield two important conclusions. First, the results suggest that the risk-adjusted returns on trading rules based on the *SI* originally could be explained by investors' surprise

about firms' earnings. Second, the diminishing relation between the *SI* and earnings announcement returns coincides with the decreasing risk-adjusted returns on *SI* portfolios discussed in Section 4.1, as well as with the heightened attention for stakeholder information in the capital market in recent years.

Table 2.6 The *SI* and earnings announcement returns: with governance indexes

We estimate the relationship between the stakeholder-relations index and cumulative earnings announcement returns using a variant of model (2.3) using cumulative abnormal return realized from 1 day before the earnings announcement date to 1 day after the announcement date. *SI* is the stakeholder-relations index, Subsample 2_t is a dummy variable that equals 1 when earnings announcements occurred during the period April 2004-December 2009 and zero otherwise. $Controls_{i,k,t-1}$ is a vector of control variables, which includes a dummy variable that captures firms' presence on KLD's list of controversial businesses, and industry fixed effects based on the 48 industry classifications from the Kenneth French Data Library. In addition, we include either the *G Index* of Gompers, Ishhi and Metrick (2003) and *G-Index**Subsample 2, or the *E Index* of Bebchuk, Cohen and Ferrell (2009) and *E Index**Subsample 2. The *t*-statistics (in parentheses) are derived from two-way clustered standard errors. The reported coefficients are multiplied by 1000 for expositional convenience. The F_1-test indicates for each regression whether the sum of the coefficients on *SI* and *SI*Subsample 2* are different from zero, and the F_2 test (Governance) indicates whether the summed coefficients on *G (E) Index* and *G (E) Index**Subsample 2 are different from zero. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively. Coefficients on the control variables other than those on the governance indexes are not reported for the sake of brevity.

SI	0.883***	0.882***	0.877***	0.882***	0.880***
	(3.23)	(3.21)	(3.17)	(3.23)	(3.22)
*SI**Subsample 2	-1.270***	-1.267***	-1.264***	-1.262***	-1.271***
	(-2.65)	(-2.65)	(-2.63)	(-2.65)	(-2.70)
Subsample 2	-0.194	-0.206	-1.361	-0.175	-1.889
	(-0.17)	(-0.18)	(-0.47)	(-0.16)	(-1.23)
G Index		-0.026	-0.096		
		(-0.19)	(-0.49)		
*G Index**Subsample 2			0.121		
			(0.42)		
E Index				-0.192	-0.596*
				(-0.67)	(-1.81)
*E Index**Subsample 2					0.713
					(1.19)
Control variables	Yes	Yes	Yes	Yes	Yes
Observations	53178	53178	53178	53178	53178
Adj. R-squared	0.002	0.002	0.002	0.002	0.002
F_1-test (β1+β2=0)	1.132	1.131	1.144	1.120	1.210
Prob. > F_1	0.29	0.29	0.29	0.29	0.27
F_2-test (Governance)			0.017		0.064
Prob. > F_2			0.90		0.80

2.4.3. Errors in analysts' forecasts

We complement our examination into errors in investors' expectations with an analysis of analysts' earnings forecasts. Although analysts' expectations do not necessarily reflect the capital market's expectations, the previous results at the very least raise the question whether analysts have misunderstood the association between stakeholder relations and firms' future earnings. Moreover, analysts have been criticized for insufficiently catering to institutional investors when it comes to integrating environmental, social, and corporate governance information in financial research (e.g. O'Loughlin and Thamotheram 2006). Therefore, if investors misunderstood the association between stakeholder relations and profitability, one could expect that analysts were at least as surprised.

In order to be consistent with the analysis of quarterly earnings announcements, we first study errors in quarterly earnings-per-share (EPS) forecasts, which we define as the difference between the actual EPS and the median forecast that I/B/E/S/ released on the closest date prior to the last day of the fiscal period. Previous studies have illustrated that inferences involving analyst forecast data are sensitive to extreme noise, skewness, outliers, and the measurement of the forecasts themselves (see, e.g., Lim (2001); Ljungqvist et al. (2009)). We address these robustness issues by analyzing alternative measures of forecast errors. Specifically, we follow the literature on analyst forecast errors and consider different ways of scaling forecast errors. We scale the errors by, respectively, the price per share at the forecast date, the assets per share, the absolute value of the median forecast and the standard deviation of the analyst forecasts. To make sure that small sample problems and outliers do not distort the median forecasts, we omit observations that either are based on forecasts from fewer than five analysts or exceed the bottom (top) 1% of the distribution.

The model we estimate takes the form:

$$FE_{i,tq} = \alpha + \beta_1 SI_{i,t-1} + \beta_2 SI_{i,t-1} Subsample2_t + \beta_3 Subsample2_t + ... \\ ... \sum_{k=1}^{K} \gamma_k Controls_{i,k,t-1} + \varepsilon_{i,tq} \quad (2.4)$$

where FE is the forecast error for quarter *q* in year *t*. As controls, we include a dummy that equals one for firms operating in a controversial industry and zero otherwise, the natural logarithm of the book to market ratio from the previous fiscal year, the natural logarithm of the market value of equity from the previous fiscal year, and industry fixed effects based on the Fama-French 48 industry definitions. In line with the previous analysis, time-variation in the relation between the *SI* and earnings forecast errors is estimated by interacting *SI* with a dummy variable that identifies forecast errors realized after March 2004.

Since earlier studies suggest that investments in stakeholder relations are mainly intangible and pay off slowly, we also study forecasts of firms' long-term earnings growth released by sell-side financial analysts in the I/B/E/S universe in order to investigate analyst forecast errors. Like Edmans (2011), we first perform pooled OLS regressions involving forecast errors defined as the long-term earnings growth that firm *i* realized at the end of fiscal year *t* minus the corresponding median value of analysts' forecasts of long-term growth made 5 years earlier (we winsorize the errors at the 1% level). Because most annual reports are filed within three months after the fiscal year-end, we measure analysts' forecasts four months after the previous fiscal year-end in order to make sure that analysts were aware of previous earnings when they made their forecast (see Core et al. (2006); Doukas et al. (2002)). Alternatively, we estimate ordered probit models after converting the earnings forecast errors to discrete variables in order to deal with the extreme noise and outliers that are common with earnings surprise data. In the probit model (Probit), the discrete variable has a value of 1 when the forecast error is greater than or equal to 10 percent, 0 when the error is between 10 percent and -10 percent, and -1 if it is equal to or below -10 percent.

According to all models of quarterly forecast errors presented in Table 2.7, firms with higher *SI* values experienced significantly higher earnings surprises over the period 1992-2004. In the subsequent years, the relationship between the *SI* and quarterly forecast errors decreased significantly under three specifications. Based on F-tests, the null hypothesis that the sum of the coefficients β_1 and β_2 is zero is not rejected in two specifications, which suggests that the *SI* is not significantly related to forecast errors in recent years. Under one other model, the relation between the *SI* and quarterly forecast errors reversed from positive to slightly negative. Indeed, it has been shown that inferences about expectational errors derived from scaled-errors in short-term analyst forecasts might be sensitive to the choice of scaling variable (see, Bebchuk et al. (2013)).

Our models of long-term forecast errors, which are presented in Table 2.8, reach a consensus. According to the OLS model, the relation between the *SI* and errors in analysts' forecast of long-term earnings growth was positive over the period 1992-2004 but close to zero and statistically not significant over the period 2004-2009. Under the ordered probit model, firms with stronger stakeholder relations were more likely to produce higher surprises in the first part of the sample period. But after the March 2004, firms with better stakeholder relations were less likely than before to have beaten analysts' long-term growth forecasts.

Furthermore, we also find no evidence that potential learning effects concerning corporate governance variables subsume the association between the *SI* and earnings surprises in models of analysts forecast errors. Table 2.9 shows that even in the presence of the *G index* and *E index*, the coefficient on the *SI* and the *SI**Subsample 2 variables are highly similar to those reported in Tables 2.8 and 2.9.[23] The coefficients on the governance variables are sensitive to the choice of earnings surprise measure, which is consistent with the results on analyst forecast errors reported in Bebchuk et al (2013).

[23] Due to space constraints, we do not report on OLS regression of errors in forecasts of long-term earnings growth. The results are available upon request.

Taken as a whole, the analyses of errors in analysts' forecasts produce results that display similarities with tests of errors in investors' expectations derived from risk-adjusted portfolio returns and abnormal earnings announcement returns.

Table 2.7 Stakeholder Index and Quarterly Errors in Analysts' Earnings Forecasts

The error in quarterly forecast is defined as the actual level of quarterly earnings minus the I/B/E/S median analyst long-term forecast closest to the error date. We report quantile (median) regressions to take the skewed distributions of the errors into account. As independent variables, we include the stakeholder-relations index (*SI*), a dummy variable (Subsample 2) that is equal to 1 whenever a forecast error is realized during the period April 2004-December 2009, an interaction term *SI**Subsample 2 that captures time variation in the relation between stakeholder relations and the dependent variable, and control variables. Sample period: April 1992 - December 2009. The *t*-statistics, derived from two-way clustered standard errors (on year-quarter and firm), are presented in parentheses. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

Variables	Percentage	Assets	Price	St. Dev
SI	1.382***	0.009***	0.006*	48.229***
	(3.99)	(3.28)	(1.90)	(7.60)
*SI**Subsample 2	-1.653***	-0.001	-0.017***	-45.071***
	(-3.42)	(-0.10)	(-3.84)	(-5.08)
Subsample 2	17.039***	0.299***	0.316***	133.819***
	(17.59)	(38.76)	(34.88)	(7.53)
Controversial business	-5.130***	-0.089***	-0.052***	-90.727***
	(-3.25)	(-7.12)	(-3.53)	(-3.14)
Log book / market equity	4.576***	-0.069***	0.090***	-102.107***
	(6.78)	(-12.78)	(14.32)	(-8.24)
Log market value of equity	-0.849**	0.008***	0.010***	97.054***
	(-2.50)	(3.01)	(3.04)	(15.57)
Constant	14.587**	-0.162***	-0.098*	-649.716***
	(2.39)	(-3.33)	(-1.72)	(-5.80)
Observations	59,320	59,320	59,320	59,320
Pseudo R-squared	0.007	0.013	0.007	0.006
F-test ($\beta1+\beta2=0$)	0.593	9.508	11.620	0.238
Prob. > F	0.44	0.00	0.00	0.63

Table 2.8 Stakeholder Index and Errors in Analysts' Forecasts of Long-Term Earnings Growth

The error in long-term growth forecast is defined as the actual five-year annualized EPS growth rate minus the I/B/E/S median analyst long-term growth forecast 56 months before the error date. We report on an OLS regression (OLS), and an ordered probit model (Probit) after we convert the forecast errors to discrete variables. In the ordered probit model, the discrete variable has a value of 1 when the forecast error is greater than or equal to 10 percent, 0 when the error is between 10 percent and -10 percent, and -1 if it is equal to or below -10 percent. As independent variables, we include the stakeholder-relations index (*SI*), a dummy variable (Subsample 2) that is equal to 1 whenever a forecast error is realized during the period April 2004- December 2009, an interaction term Stakeholder*Subsample 2 that captures time variation in the relation between stakeholder relations and dependent variable, and control variables. Sample period: April 1992-December 2009. The t-statistics (*z*-statistics) in parentheses are derived from two-way clustered standard errors (on year and firm). *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

	OLS	Probit
SI	0.267**	0.012*
	(1.97)	(1.69)
*SI**Subsample 2	-0.388	-0.026**
	(-1.55)	(-1.96)
Subsample 2	3.647	0.206*
	(1.54)	(1.70)
Controversial business	0.442	-0.011
	(0.42)	(-0.21)
Log book / market equity	-5.156***	-0.286***
	(-8.11)	(-8.58)
Log market value of equity	2.071***	0.104***
	(10.90)	(7.19)
Constant	-29.576***	
	(-13.10)	
Observations	15,191	15,191
Adj ./Pseudo -R-squared	0.080	0.043
F test / Chi-square test (β1+β2=0)	0.362	1.929
Prob. > F	0.55	0.17

Table 2.9 The *SI* and earnings forecast errors: controlling for governance indexes

For table description see table 2.7 for the OLS results using quarterly forecast errors en table 2.8 for the ordered probit models that analyze long term forecast errors. In all analyses we add governance indexes (and an interaction with Subsample 2) to the set of controls.

	OLS								Ordered probit	
	Percentage	Percentage	Assets	Assets	Price	Price	St. Dev.	St. Dev.	Long-term	Long-term
SI	1.570***	1.667***	0.011***	0.011***	0.009***	0.008***	53.688***	51.022***	0.012*	0.012*
	(5.30)	(5.77)	(4.21)	(4.37)	(3.51)	(3.34)	(7.29)	(6.52)	(1.70)	(1.70)
*SI**Subsample 2	-1.437***	-1.586***	0.002	0.002	-0.013***	-0.014***	-47.109***	-45.032***	-0.024	-0.024
	(-3.35)	(-3.79)	(0.64)	(0.67)	(-3.70)	(-3.78)	(-4.42)	(-3.97)	(-1.55)	(-1.60)
G index	-0.346		-0.002		-0.001		-18.487***		0.006	
	(-1.47)		(-0.77)		(-0.47)		(-3.16)		(0.70)	
*G index**Subsample 2	-0.322		-0.015***		-0.004		16.043**		0.002	
	(-0.99)		(-5.23)		(-1.31)		(1.98)		(0.16)	
E index		-1.252***		-0.006		-0.003		-43.972***		0.024
		(-2.74)		(-1.41)		(-0.71)		(-3.55)		(1.47)
*E index**Subsample 2		0.344		-0.027***		-0.004		27.083		0.003
		(0.55)		(-4.94)		(-0.70)		(1.59)		(0.17)
Subsample 2	20.203***	16.776***	0.437***	0.364***	0.344***	0.322***	7.892	106.596**	0.168	0.171
	(6.25)	(9.71)	(15.09)	(24.12)	(12.85)	(21.21)	(0.10)	(2.28)	(1.10)	(1.10)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	43,554	43,554	43,554	43,554	43,554	43,554	43,554	43,554	12,197	12,197
Pseudo R-squared	0.006	0.006	0.013	0.013	0.007	0.007	0.007	0.007	0.040	0.041
F_1-test ($\beta1+\beta2=0$)	0.169	0.065	22.041	23.770	2.814	3.791	0.668	0.488	0.871	0.991
$P > F_1$	0.68	0.80	0.00	0.00	0.09	0.05	0.41	0.49	0.35	0.32
F_2-test (Governance)	8.429	4.047	67.101	68.913	5.435	2.882	0.182	1.906	1.023	3.543
$P > F_2$	0.00	0.04	0.00	0.00	0.02	0.09	0.67	0.17	0.31	0.06

2.5. Additional tests

2.5.1. Alternative factor models

Up to this point, our estimates of average risk-adjusted return on top- and bottom-ranked portfolios have been derived from the Carhart (1997) four-factor model. To ensure that the observed decreasing risk-adjusted returns documented in Section 4.1 are not an artifact of that specific model, we report the estimates of risk-adjusted returns that we obtain under alterative specifications in this section.

The first two models we report on in Table 2.10 are a one-factor model, and the three-factor model of Fama and French (1993). The third model presented in Table 2.9 is an alternative four-factor model brought forward by Cremers, Petajisto and Zitzewitz (2013), who argue that a model containing the momentum factor augmented with market, size and value premiums based on tradable indexes better captures returns than the standard multifactor models from Fama and French (1993) and Carhart (1997). The last model we use to measure risk-adjusted returns is the Carhart (1997) model extended with the traded liquidity factor of Pastor and Stambaugh (2003).

Taken together, the intercepts from these alternative factor regressions reinforce the idea that positive risk-adjusted returns associated with the *SI* eventually ceased to exist. Independent of the factor model, the difference in average risk-adjusted return between top-ranked and bottom-ranked portfolios was positive, economically large, and significant at the conventional cut-off levels during the period 1992-2004. For the period April 2004-December 2009, none of the factor models produces a risk-adjusted return that is significantly different from zero.

Table 2.10 Performance under alternative factor model specifications

Every year, starting in April 1992, we rank stocks based on the stakeholder-relations index (*SI*) and assign the top (bottom) third, fourth, or fifth of all ranked stocks to a top-ranked (bottom-ranked) portfolio. Using alternative factor models, we then estimate differences in risk-adjusted portfolio returns between top and bottom ranked portfolios over the period April 1992-December 2009 and the subperiods April 1992-March 2004 and April 2004–December 2009. The factor models we consider for performance evaluation are, respectively, a 1-factor model that includes as explanatory variable the CRSP value-weighted return described in equation (2.1), the three-factor model of Fama and French (1993), the four factors proposed by Cremers, Petajisto, and Zitzewitz (2012), and the Carhart (1997) four-factor model augmented with the Pastor and Stambaugh (2003) liquidity factor (Pastor-Stambaugh). Reported are annualized risk-adjusted returns for equal-weighted portfolios, with *t* statistics in parentheses. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

Factor model	1-factor	Fama-French	Cremers et al.	Pastor-Stambaugh
Top minus bottom third				
1992-2009	2.10%	2.29%*	1.73%	1.61%
	(1.56)	(1.72)	(1.31)	(1.22)
1992-2004	3.88%**	5.38%***	2.83%*	3.49%**
	(2.18)	(3.35)	(1.94)	(2.41)
2004-2009	-1.40%	-1.90%	-1.43%	-2.06%
	(-0.73)	(-1.17)	(-0.89)	(-1.39)
Top minus bottom fourth				
1992-2009	3.48%***	3.89%***	3.37%***	3.43%***
	(2.78)	(3.18)	(2.68)	(2.70)
1992-2004	4.76%***	6.36%***	4.38%**	5.22%***
	(2.76)	(3.86)	(2.60)	(3.06)
2004-2009	0.99%	0.86%	0.92%	0.79%
	(0.68)	(0.60)	(0.64)	(0.55)
Top minus bottom fifth				
1992-2009	3.44%**	3.62%**	3.04%**	2.81%*
	(2.31)	(2.52)	(2.00)	(1.93)
1992-2004	4.69%**	6.32%***	3.56%**	4.33%**
	(2.35)	(3.49)	(2.16)	(2.58)
2004-2009	1.51%	1.36%	1.46%	1.24%
	(0.82)	(0.74)	(0.80)	(0.69)

2.5.2. Alternative firm scores based on KLD data

This section examines alternative firm scores derived from KLD (henceforth, *Alternative KLD*) data that have been associated with positive abnormal returns in the literature. In particular, we pay attention to measures used by Kempf and Osthoff (2007) who document positive risk-adjusted returns associated with several different stock ranking approaches based on a composite of indicators from KLD. They use indicators from six KLD categories: community, diversity, employee relations, environment, human rights, and product. To get an overall score for each firm called "Combination 1", they transform the concerns by taking the binary complements, then sum up the scores from all KLD criteria, and normalize this sum so that the score ranges from zero to one. In addition, they also create an overall score that is first subject to a so-called "negative screen" ("Combination 2"), by excluding all firms that are involved in at least one of the controversial business areas that are identified by KLD. We also consider a "Best-in-Class" version of their Combination 1, where we rank firms on Combination 1 relative to the industry average Combination 1 scores (using the Fama-French 10 industry classifications). Our fourth alternative measure is obtained by simply taking the sum of all strengths that KLD identified for a firm in a given year minus the sum of all concerns that KLD identified across all possible indicators ("Strengths – concerns").

Panel A in Table 2.11, shows that all of these alternative measures for ranking stocks lead to top-minus-bottom third portfolios that produced positive risk-adjusted returns during the period April 1992-March 2004, but which did not deliver positive risk-adjusted returns during the remainder of the sample period. In Panel B, we show the results of replacing the *SI* by, respectively, Combination 1, Combination 2, and Strengths-Concerns measures in regressions involving earning announcement returns (measured from 1 day before to 1 day

after each announcement).[24] Consistent with diminishing errors in expectations, the coefficient estimates suggest that all three alternative measures were positively related to earnings announcement returns prior to April 2004 but not in the period thereafter.

Table 2.12 reports on the inclusion of the *Alternative KLD* measures in the different models of quarterly and long-term analyst forecast errors. Although the results are more mixed than those from portfolio analyses and earnings announcement regressions, the coefficients in twelve out of the fifteen specifications in Table 2.12 suggest that these alternative measures related positively to forecast errors during the period 1992-2004. In nine specifications, the coefficients on the interaction term *Alternative KLD*Subsample 2* suggests that the relation decreased significantly during the period April 2004-December 2009.

2.5.3. Stakeholder relations and future profitability

For investors to overlook the difference in future profits between top- and bottom-ranked firms, it is important to verify that an association between the *SI* and future profitability exists to begin with. For this reason, we also show results of regressing firms' future operating performance, as measured by return on assets, on the lagged *SI* and a set of control variables:

$$ROA_{i,t} = \alpha + \beta_1 SI_{i,t-1} + \sum_{k=1}^{K} \gamma_k Controls_{i,k,t-1} + \varepsilon_{i,t} \quad (2.5)$$

where $ROA_{i,t}$ is the accounting return on assets (defined as either operating income after depreciation and amortization divided by total assets, or net income divided by assets) for the fiscal year subsequent to the year for which KLD reports its information; and $Controls_{it-1}$ is a vector of control variables. The vector of control variables includes a dummy for

[24] The Combination 2 measure is not included in the earnings announcement regressions since a controversial business indicator variable is separately included in the models.

controversial industries, the natural logarithm of the book-to-market ratio, the natural logarithm of total assets, the natural logarithm of firm age identified as the number of months the firm first appeared in the CRSP returns database until December of the year, a dummy for Delaware incorporation, R&D divided by total sales, capital expenses divided by total assets in conjunction with dummy variables that identify non-reported R&D and capital expenses, and year- and industry-fixed effects (also see Jiao 2010)). All variables that are not reported as a natural logarithm are winsorized at the 1% level to account for outliers.[25] These variables (except firm age) are constructed using data from Compustat.

Table 2.13 shows the coefficients from the regressions together with *t*-statistics derived from two-way clustered standard errors. The coefficients on the control variables have signs that are consistent with the majority of studies on the determinants of profitability. Most important to this study is the coefficient on the *SI*. Independent of the model employed, we find that the relation between the *SI* and ROA is positive and statistically significant at the conventional significance levels. Hence, these results suggest that information about corporate stakeholder relations is relevant in understanding firms' future profits.

[25] Winsorizing or trimming at different levels does not qualitatively alter our results.

Table 2.11 Risk-adjusted portfolio returns and earnings announcement returns: alternative measure of stakeholder relations

In Panel A, we report on top-ranked and bottom-ranked portfolios that are formed based on alternative firm-level measures derived from the KLD database. Starting in April 1992, we rank stocks based one of four alternative measures based on KLD indicators and assign the top (bottom) third of all ranked stocks to an top-ranked (bottom-ranked) portfolio. Using the Carhart (1997) four-factor model, we then estimate differences in risk-adjusted portfolio returns between top- and bottom-ranked portfolios over the period April 1992-December 2009 and the subperiods April 1992-March 2004 and April 2004–December 2009. The alternative measures are, respectively, Kempf and Osthoff's (KO, 2007) Combination 1, Combination 2, and Best-in-class measures, and an industry-unadjusted version of the *SI* (which thus simply aggregates all strengths and subtracts all concerns reported by KLD). Reported are annualized risk-adjusted returns for equal-weighted portfolios, with *t* statistics in parentheses. In Panel B we report on estimating three-day earnings announcement returns (-1,+1) using model (2.3) after replacing the *SI* by one of four alternative measures that are based on KLD indicators. (*Alternative KLD*). The F-test and corresponding p-value indicate for each regression whether the sum of the coefficients on the *Alternative KLD* measure and *Alternative KLD**Subsample 2 are different from zero. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

Panel A: Risk-adjusted portfolio returns, top minus bottom third

	Carhart (1997) four-factor alpha		
Cut off	1992-2009	92-04	04-09
KO Combination 1	1.14%	2.88%**	-2.30%
	(1.03)	(2.38)	(-1.53)
KO Combination 2	1.30%	2.78%**	-2.04%
	(1.11)	(2.07)	(-1.25)
KO Best-in-class	1.44%	2.36%**	0.07%
	(1.35)	(2.00)	(0.05)
Strenghts - Concerns	1.94%	4.27%***	-3.01%*
	(1.54)	(2.67)	(-1.86)

Panel B: Earnings announcement returns (-1,+1)

	KO Combination 1	KO Best-in-class	Strenghts – concerns
Alternative KLD	19.378***	22.713***	0.843***
	(2.70)	(2.88)	(3.07)
*Alternative KLD**Subsample 2	-27.527***	-28.215**	-1.070***
	(-2.63)	(-2.29)	(-2.91)
Subsample 2	18.819**	-0.378	-0.062
	(2.45)	(-0.33)	(-0.05)
Controversial business	2.564**	2.574**	2.298**
	(2.37)	(2.38)	(2.06)
Constant	-3.924	10.222***	10.078***
	(-0.65)	(3.04)	(2.83)
Observations	91,290	91,290	78,340
Adj. R-squared	0.002	0.002	0.002
F-test ($\beta1+\beta2=0$)	1.030	0.336	0.681
Prob. > F	0.31	0.56	0.41

Table 2.12 Analysts forecast errors and alternative measures of stakeholder relations

The error in quarterly forecast is defined as the actual level of quarterly earnings minus the I/B/E/S median analyst long-term forecast closest to the error date, scaled by either the absolute value of the median forecast (*Percentage*), assets per share (*Assets*), price per share (*Price*), or the standard deviation of analysts' forecasts (*St. Dev.*). The error in long-term growth forecast is defined as the actual five-year annualized EPS growth rate minus the I/B/E/S median analyst long-term growth forecast 56 months before the error date. We estimate models of quarterly and long-term forecast errors, using alternative measures based on KLD and a set of control variables as explanatory variables. The alternative measures (*Alternative KLD)* are, respectively, Kempf and Osthoff's (2007) Combination 1 and Best-in-class measures, and an industry-unadjusted version of the *SI* (i.e., the sum of all strengths minus the sum of all concerns reported by KLD). Coefficients on the control variables are not reported due to space constraints.

Forecast error	*Alternative KLD*	*Alternative KLD**Subsample 2	Subsample 2	Observations	Pseudo R2	F-test (β1+β2=0)	Prob. > F
Percentage							
Combination 1	34.800***	-4.070	22.193***	66843	0.007	17.139	0.00
Best-in-Class	46.821***	-29.051**	17.177***	66843	0.007	3.719	0.05
Strenghts - concerns	0.975***	-1.000**	17.631***	59320	0.007	0.008	0.93
Assets							
Combination 1	0.228***	-0.053	0.350***	66843	0.013	9.541	0.00
Best-in-Class	0.287***	0.132	0.292***	66843	0.013	34.430	0.00
Strenghts - concerns	0.005**	0.010**	0.303***	59320	0.012	24.490	0.00
Price							
Combination 1	0.136*	-0.229**	0.465***	66843	0.007	1.858	0.17
Best-in-Class	0.259***	-0.085	0.301***	66843	0.007	4.487	0.03
Strenghts - concerns	0.003	-0.013***	0.319***	59320	0.007	9.265	0.00
St. Dev.							
Combination 1	834.612***	635.355***	-205.667	66843	0.006	82.563	0.00
Best-in-Class	1,025.933***	-639.855**	126.001***	66843	0.007	3.917	0.05
Strenghts - concerns	30.467***	-23.548**	158.035***	59320	0.007	1.012	0.31
Long-term							
Combination 1	-5.411	-43.063***	30.451***	17,347	0.101	38.972	0.00
Best-in-Class	0.320	-16.832**	3.346	17,347	0.086	8.432	0.00
Strenghts - concerns	0.495***	-0.692**	4.042*	15,191	0.081	0.862	0.35

Table 2.13 Stakeholder relations and profitability

This table reports on pooled regressions with accounting return on assets (ROA) as dependent variable and the *SI* in conjunction with control variables as independent variables as in model (2.5). Return on assets (ROA) is defined as either the ratio of operating income (after depreciation and amortization) divided by total assets or net income divided by total assets. The control variables include a dummy variable capturing firms' controversial business involvement (alcohol, gambling, firearms, military, nuclear power, tobacco) according to KLD, the logarithm of the book-to-market ratio, the logarithm of total assets, R&D expenses scaled by sales, capital expenditures scaled by total assets, dummy variables that identify non-reported R&D and capital expenditures, and year fixed-effects, and industry-fixed effects based on 48 industry classifications from the Kenneth French Data Library. The *t*-statistics (in parentheses) are derived from two-way clustered standard errors. Sample period 1992-2009. *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.

	Operating income / assets	Net income / assets
SI	0.004***	0.004***
	(4.81)	(4.01)
Controversial business	-0.003	-0.005
	(-0.87)	(-1.19)
Log book / market equity	-0.025***	-0.011***
	(-6.60)	(-2.98)
Log total assets	0.006***	0.005***
	(3.36)	(3.65)
Log age	0.006***	0.005**
	(4.41)	(2.33)
Delaware	-0.009***	-0.012***
	(-3.63)	(-4.86)
CAPEX / assets	0.030*	0.010
	(1.76)	(0.68)
R&D / sales	-0.083***	-0.060***
	(-21.73)	(-13.75)
R&D Dummy	0.012***	0.011***
	(3.35)	(3.05)
CAPEX / assets dummy	-0.001	-0.000
	(-0.34)	(-0.05)
Constant	-0.001	-0.055**
	(-0.07)	(-2.19)
Observations	21,310	20,643
Adj. R-squared	0.348	0.233
Year FE	YES	YES
Industry FE	YES	YES

2.6. Conclusion

Many investors justify the integration of stakeholder information – nowadays under the heading of "ESG" information – in portfolio selection by the view that corporate stakeholder relations are associated with (intangible) value in a manner that is not fully understood by the financial market. Although this view is not necessarily counterintuitive in the short run, investors' public hunt for "mispriced" information that generates superior risk-adjusted returns eventually comes as a double-edged sword. Economic logic teaches us that increased attention to value-relevant information makes potential "mispricing" short-lived.

This paper shows that trading strategies that use a stakeholder-relations index generated risk-adjusted returns that were economically and statistically significant over the period 1992-2004, but that were largely non-significant over the period 2004-2009. This finding is in line with our premise that a stakeholder-relations index predicted risk-adjusted returns due to errors in investors' expectations, but ultimately ceased to do so as attention for stakeholder issues increases.

Our findings are based on three complementary approaches, commonly used in empirical studies on stock market anomalies. A portfolio approach, an event study around quarterly earnings announcements, and an analysis of errors in analysts' forecasts all point in the same direction, and show that errors in expectations that arise due to difficulties in assessing the value of stakeholder relations investments, are not persistent.

Furthermore, the paper suggests, using a statistical procedure described in Quandt (1960), that a break in the analysis occurred around the year 2004. This seems in line with annual statistics on the number of shareholder proposals on stakeholder issues, and with previous studies that document a strong increase in the number of CSR reports published by companies.

The implications of our findings are that those institutional investors that pursue both financial and social goals have empirical foundations for integrating stakeholder issues in investment decisions. However, the contribution of stakeholder information to generating abnormal returns does not persist in the long term. Our findings also imply that companies should place stakeholder issues higher on the corporate agenda given that stakeholder management nowadays appears to be more fully appreciated by investors.

Chapter 3

3. Can investors profit from social tastes? Evidence from mutual fund holdings

3.1. Introduction

Mutual fund attributes are important for investors as they use these attributes to select and evaluate investment products. Well-known examples are past return, investment style, risk profile, fund family brand, fund size, and fund age (e.g. Sirri and Tufano 1998). Another attribute bundle that is increasingly receiving attention is the social responsibility of the mutual fund (e.g. Bollen 2007, Renneboog et al. 2011). To investigate this attribute bundle, researchers have compared funds with a socially responsible investment (henceforth SRI) label to funds without one (henceforth conventional funds). The consentient findings show that SRI and conventional funds have earned similar risk-adjusted returns (see Derwall, Koedijk, and Ter Horst (2011) for a review), though, fund flows can be affected by an SRI label (Bollen 2007, Benson and Humphrey 2008, and Renneboog et al. 2011).[26] In this paper we differentiate on two aspects by evaluating multiple social responsibility attributes for all U.S. equity mutual funds. Holdings-data enables us to identify exposures to several social responsibility attributes. Using this approach, we provide novel insights into funds' exposures to social controversies and social responsibility as well as performance aspects of these attributes.

[26] Benson and Humphrey (2008) and Renneboog et al. (2011) find that SRI investors are less sensitive to fund flows. Whereas Bollen (2007) reports that investors are more sensitive to past positive returns and there is no difference when considering past negative returns.

The mutual fund literature up to date has compared SRI and conventional funds as if they differ substantially on several social responsibility characteristics. However, we take a different approach to identify a fund's social responsibility by looking at exposures to several company-characteristics associated with SRI. This method enables us to take into account that social responsibility is a multi-facetted concept meaning that mutual funds can be considered socially responsible on one area and irresponsible on another. In addition, we explore the possibility that conventional fund holdings are affected by managers' tastes for companies with social responsibility characteristics. Papers that have compared holdings of SRI to conventional funds have reported that on average SRI funds have lower exposure to stocks that are deemed socially irresponsible (e.g. Di Giuli and Kostovetsky 2014, Kempf and Osthoff 2011) and that these funds have higher exposures to stocks of firms with high scores from social responsibility rating agencies, so called environmental, social, and governance (henceforth ESG) scores (Kempf and Osthoff 2011) on average.

Social responsibility attributes can also affect mutual fund returns through tastes for or against assets. A growing body of studies suggests that societal norms can influence the capital market due to investors' tastes against (for) assets of companies they deem socially objectionable (desirable). The general prediction is that tastes can affect asset prices because investors see the assets as consumption goods (Fama and French 2007). In other words, attributes other than the payoff structure can matter to investors. More specifically, asset prices can be affected when a significant number of investors have similar tastes (Fama and French 2007). In this paper we consider social responsibility attributes that could matter to investors through societal norms. Common examples of such social responsibility attributes are companies that earn revenues from the tobacco, alcohol, gambling, and weapons industries (e.g. Hong and Kacperczyk 2009), companies with relatively good corporate social

responsibility practices (El Ghoul, Guedhami, Kwok, and Mishra 2011) or stocks of heavily polluting firms (Heinkel, Kraus, and Zechner 2001)[27].

Despite the notion of a "social premium" in academic studies, it remains an open empirical question whether U.S. investors can profit from social norms by investing in stocks that are sensitive to social tastes through the lens of a significant number of investors. On the one hand, evidence teaches us that socially controversial stocks have either higher expected returns as implied by stock prices or earned higher realized returns than socially acceptable stocks (e.g. Fabozzi, Ma, and Oliphant 2008, Hong and Kacperczyk 2009, El Ghoul et al. 2011, Derwall et al. 2011, Chava 2013). On the other hand, there is overwhelming evidence that so-called socially responsible (i.e. SRI-labeled) mutual funds, which explicitly proclaim to screen out controversial assets, have not underperformed conventional mutual funds over various investment horizons; see e.g., Derwall al. (2011) for a review. This evidence could be taken to imply that the effects of social tastes on stock prices have no meaningful investment implications once trading costs, benchmark constraints, and illiquidity are accounted for.

In this paper, we aim to determine the true economic significance of social investment decisions by studying the actual holdings of U.S. equity mutual funds over the period January 2004 to December 2012. For the entire universe of mainly domestic U.S. equity funds, we transform mutual fund holdings to scores that measure a fund's exposures to oft-cited socially sensitive assets: stocks from tobacco, alcohol, and gambling (Hong and Kacperczyk 2009), weapons manufacturers and nuclear operations (Statman and Glushkov 2009, Derwall et al. 2011), and firms with ESG concerns or strengths (e.g. El Ghoul et al. 2011, or specifically environmental concerns and strengths; Chava 2013). Using these scores, we estimate the payoff that mutual funds in reality enjoy for every fraction of wealth that is invested in socially sensitive stocks. We are able to contrast these estimates with the returns of

[27] Other related theoretical studies include Angel and Rivoli (1997) and Gollier and Pouget (2012).

hypothetical portfolios that are formed by ranking stocks on their social responsibility attributes, and with the conclusions reached by studies on SRI mutual funds.

We document an economically and statistically significant payoff associated with mutual funds' exposure to stocks from tobacco, alcohol and gambling sectors, often referred to as "sin stocks" (Hong and Kacperczyk 2009). Pooled regression estimates indicate that a 100 percentage-point increase in the fraction of wealth invested in sin stocks is associated with a higher risk-adjusted return of about 4% to 5% on an annualized basis, which is roughly similar to that suggested by analyses of hypothetical sin stock portfolios. We do not find statistically significant payoffs associated with funds' exposures to the other types of controversies. In addition we find some evidence that exposure to firms with very strong social responsibility profiles has a negative effect on performance.

While these performance effects are statistically significant, we find that most mutual funds do not display full exposures to the socially sensitive assets we consider. Due to limited cross-sectional variation in socially sensitive investment exposures, the annualized risk-adjusted return spread between a portfolio of funds with highest fund scores and the lowest-ranked counterparts was not significant for all measures over the period January 2004 – December 2012.

We also use the same logic to explain why SRI-labeled funds have not performed differently from non-SRI funds (see Derwall et al. (2011) for a review). We show that neither the "SRI" label nor the type of screen that SRI funds disclose in prospectuses and other public sources is adequate for distinguishing mutual funds along their socially sensitive investments like sin stocks. A significant part of conventional funds in our sample have significantly lower exposures to sin stocks (30%), a broader definition of sin stocks that includes firms in the weapons/defense and nuclear energy industries (10%) and firms with ESG concerns (10%) score of the average SRI fund.

These findings make several contributions to the literature. First, we contribute to the literature on social tastes in markets. Our results suggest that the effects of tastes have a material impact on mutual fund holdings of both socially responsible investment funds and conventional funds. The advantage of studying mutual fund holdings is that these funds have traded stocks based on real prices and their returns are generated in the presence of real-time transaction costs and trading restrictions.

Second, we contribute to the literature on socially responsible investing. Our results make a case for studying the effects of social tastes on performance based on the actual holdings of mutual funds. Although it has been common practice to compare explicit SRI funds and their stated social investment screens with conventional funds in order to determine the influence of social investing on performance, we show that such a comparison masks the true role of social dimensions in explaining investment returns. This evidence supports the view that the effects of social tastes on investments reach beyond explicitly SRI labeled investments.

Third, we contribute to the literature that uses mutual funds to assess the economic significance of return predictor variables that are uncovered in studies on the cross-section of stock returns. Previous studies that use mutual funds have shown that common stock anomalies such as size, value, momentum, and accruals effects in stock returns are different on paper than in reality (e.g. Ali et al. 2008; Huij and Verbeek 2009). We find that exposure to sin stocks and exposure to firms with good social responsibility profiles (meaning few ESG concerns and many ESG strengths) incrementally explains variation in mutual fund performance.

The next section of this Chapter describes the data we use to identify socially controversial stocks, the exposures to socially sensitive stocks of U.S. mutual funds, and financial data on the stocks, mutual funds, and benchmark portfolios that are central to this

study. Section 3.3 outlines the computation of mutual funds controversial investment scores and presents initial evidence on social tastes in investment portfolios. In Section 3.4 we analyze the effects of social tastes on investment performance. Section 3.5 concludes this study.

3.2. Data

Using the CRSP Mutual Funds Database, we gather information about the holdings, monthly returns, and characteristics of mainly domestic U.S. equity mutual funds from January 2004 up to December 2012. We exclude funds that had less than 75 percent of their assets invested in U.S. equities, index funds, specialty funds, global funds, micro cap funds, and ETFs.[28] In order to identify which of the securities held by the funds are socially controversial, we match all their equity holdings with the MSCI STATS database. The MSCI STATS database (formerly known as Kinder, Lydenberg, and Domini (KLD)) has since 2003 provided, on an annual basis, more than 50 indicators from 7 broader "Environmental, Social, and Governance" (ESG) categories covered for all constituents of the Russell 3000 universe. Among the indicators that STATS covers are controversial business indicators that span a firm's involvement in tobacco, alcohol, gambling, firearms, military, and nuclear power, as well as "concerns" and "strengths" indicators in ESG areas beyond the aforementioned controversial businesses. We require at least 25 stocks *and* 75 percent of the equities of each fund to have a successful match with STATS. We also match the firms covered in STATS with the CRSP stocks database in order to obtain monthly stock returns and market capitalization at the security level from January 2004 to December 2012. Furthermore, we measure risk-adjusted returns of mutual funds and stocks by using Cahart's (1997) four-factor

[28] We keep funds with the following investment objectives (retained from Lipper data as well as Fund names): Capital Appreciation, Growth, Growth Income, Income, Mid Cap, Small Cap.

model, with the factor data taken from the Kenneth French Data Library.[29] Finally, we require the Carhart (1997) asset-pricing model to explain at least 50% of the variation in the mutual fund returns. This requirement serves as the gatekeeper to ensure we are using only U.S. domestic equity funds. The selection procedure results in a sample of 6647 mutual fund-year observations that represent the degree of controversial investments from January 2004 to December 2012. Considering that we initially rank 17504 fund-year observations from the holdings files we exclude a big part of the sample. However, we take comfort in the fact that our final sample covers over 89% of the total market capitalization of all STATS ranked fund equities. This implies that we drop very small funds and funds with few equity investments under management.

Among the set of mutual funds in the U.S. are certain funds that explicitly state the use of social screens in investment decisions. These "so-called" socially responsible mutual funds (SRI funds) have traditions of adopting "negative screens" to their investable stock universe, which involves the exclusion of stocks that are inconsistent with either societal norms or with the personal, religious, or political values of the funds' clientele. In order to determine which of the mutual funds in our sample are explicit SRI funds, we use Morningstar Premium, the U.S. Social Investment Forum, SocialFunds.com, and previous studies on SRI mutual funds. To determine the accuracy of these sources, we hand-collected information about the social responsibility screens that the funds claim to apply in their investments, using the funds' websites, prospectuses, and occasionally email correspondence with fund managers. A fund is confirmed to be explicitly socially responsible (SRI = 1) if the fund indicates that it applies at least one of the screens that we consider to the investment

[29] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#HistBenchmarks, these factors are based on Fama French (1993) extended with the momentum factor from Carhart (1997).

opportunity set.[30] The number of U.S. SRI equity funds with confirmed investment screens increases over time, from 52 in 2004 to 72 in 2012.

3.3. Empirical analysis: Socially sensitive investments

3.3.1. Measuring mutual funds' SR investment exposure

For every mutual fund in our sample, we determine their scores quarterly along four dimensions of stock investment associated with investor tastes. The first dimension covers "sin stocks", defined as the stocks of companies that according to STATS earn revenues from the sectors tobacco, alcohol and gambling. According to Hong and Kacperczyk (2009), especially sin stocks are shunned by investors because of societal norms against funding "vice". We also consider a broader set of "sin stocks", which additionally includes stocks of companies that are involved either in the firearms and military industry or with nuclear operations (see, e.g., Statman and Glushkov 2009, Derwall et al. 2011). In addition to these controversial business areas, we also consider ESG indicators since investors' tastes for assets can be influenced by the corporate social responsibility (CSR) of firms (Fama French 2007). Both theoretical and empirical studies suggest that environmental concerns may be associated with higher expected returns due to norms against these assets (e.g. Heinkel et al. 2000, Chava 2013). In addition, El Ghoul et al. (2011) find that a broader set of CSR indicators are related to firms' cost of capital. Firms with better (worse) ESG scores have a lower (higher) cost of equity capital, which is in line with the view that investor tastes are relevant for firms financing costs and thus for investors.

[30] We verify the responsible investment screens applied by the funds in this set on the presence of screens concerning alcohol, gambling, tobacco, weapons, and nuclear operations. In addition to these screens we evaluate the presence environmental, social, and governance screens, and other "social" screens (community, diversity, employee, environment, human rights, and governance).

To derive the four scores for mutual fund *i*, we use the fund's quarterly holdings and value weight the firm level *j*. Using the controversial business indicators and ESG indicators from STATS, we arrive at the following mutual fund scores.

$$FundSIN_{i,yr,q} = \sum_{j=1}^{J} weight_{i,j,yr,q} * Dfirm\_sin_{j,yr-1} \quad (3.1a)$$

$$FundBROADSIN_{i,yr,q} = \sum_{j=1}^{J} weight_{i,j,yr,q} * Dfirm\_broadsin_{j,yr-1} \quad (3.1b)$$

$$FundCON_{i,yr,q} = \sum_{j=1}^{J} weight_{i,j,yr,q} * Adj\_firm\_CON_{j,yr-1} \quad (3.1c)$$

$$FundSTR_{i,yr,q} = \sum_{j=1}^{J} weight_{i,j,yr,q} * Adj\_firm\_STR_{j,yr-1} \quad (3.1d)$$

where $Dfirm\_sin$ is a dummy variable that equals 1 if stock j held by the fund is associated with tobacco, alcohol, or gambling sectors according to STATS, $Dfirm\_broadsin$ is a dummy variable that equals 1 if stock j is associated with tobacco, alcohol, gambling, firearms and military, or nuclear operations, and $weight$ measures the fraction of Russell 3000 total net assets under management that fund *i* is invested in stock *j*. We also created two measures of a fund's exposure to ESG concerns and strengths using all ESG indicators from STATS adjusted for industry and market capitalization since these are known to affect the number of ESG indicators (e.g. Kempf and Osthoff 2007, Statman and Glushkov 2009). For example, large firms in the natural resources industry have more ESG concern *and* strength indicators than do small and midsized financial services firms. The adjustment is done through year-by-year OLS regressions of the total number of ESG concerns (strengths) on market capitalization, squared market capitalization, and Fama French 48 (minus one) industry dummies. We predict a firms ESG concerns (strengths) using the coefficients of

these regressions forming E(firm_CON) and E(firm_STR).[31] Subsequently we calculate the adjusted ESG measures as follows:

$$Adj\_firm\_CON_{j,yr} = firm\_CON_{j,yr} - E(firm\_CON)_{j,yr} \quad (3.2a)$$

$$Adj\_firm\_STR_{j,yr} = firm\_STR_{j,yr} - E(firm\_STR)_{j,yr} \quad (3.2b)$$

These adjustments give us a measure for ESG performance relative to industry and size peers. Without these adjustments we would be picking up firm size and industry effects in both the concern and strengths measures. Because we observe that the quarterly holdings data is not complete for all funds (funds tend to report only (semi-)annually especially in the earlier years of our sample) we use the yearly average of all quarterly available fund scores for each mutual fund and retain all funds that have at least two quarterly scores available.

Descriptive statistics on the mutual fund score-year observations from 2004 up to 2012 are presented in Table 3.1, along with other characteristics of the mutual funds in our sample. Table 3.1 shows that the average fund has 4.1 percent of all assets under management invested in sin stocks (mean *FundSIN* = 0.041) and 13.4 percent invested in a broader set that additionally covers firearms and military and nuclear operations (mean *FundBROADSIN* = 0.134). Furthermore, 2.6 percent of all fund-year observations trace back to explicitly SRI-labeled funds. A further breakdown of these SRI funds by the types of social investment screens they claim to employ indicates that almost all of the funds state some form of sin stock screening, with 2.1 percent employing at least one screen within the "narrow" definition of sin stocks (SRI SIN screen = 0.021), and 2.2 percent employing at least one screen within the broader definition that also include weapons and nuclear operations (SRI

[31] Each year we estimate: $firm\_SCORE_j = \alpha + \sum_{k=1}^{48-1} \beta_k * Dindustry_{k,j} + \gamma * MCap_j + \delta * MCap_j^2 + \varepsilon_j$
Where firm_SCORE is either STR or CON. Subsequently we derive $E(firm\_CON)_j$ from the regression coefficients.

BROADSIN screen = 0.022). Explicit ESG portfolio screening is a little less common as it is done in 1.7% (SRI ESG screen = 0.017) of all fund-years.

Table 3.1 also shows that the average domestic equity fund is 178 months old, has 1.6 billion of assets under management, and charges 1% of expenses excluding 12b1 fees. Of all funds in our universe, 52.1 percent of the fund-year observations correspond to funds that have at least one class with load fees.

The histograms A to D of annual mutual fund scores reported in Figure 3.1 indicate that U.S. mutual funds are to a varying degree invested in controversial stocks, with a few funds having more than 50 percent of their total assets under management invested in controversial firms. Among those that score high on *FundSIN* and *FundBROADSIN* are the well-known VICE Fund, Fidelity's Defense & Aerospace Portfolio, and several Industrial funds. On the other hand, a significant number of mutual funds had no capital invested in companies that STATS associates with Tobacco, Alcohol, Gambling, Weapons and Military, and Nuclear operations related businesses (Figures 3.1A and 3.1B).[32] A closer look at *FundCON* shows that at least some funds with ESG screens follow through on their prospectus. Amongst the group with the lowest *FundCON* (meaning lowest exposure to ESG concerns) we find fund-years from the following SRI funds: Calvert, Parnassus, and Professionally Managed Funds: The Women's Equity Fund. However within the group with the lowest *FundCON* we also observe numerous growth equity funds like the John Hancock Growth Trends Fund and the Barrett Growth Fund. These observations hint in the direction of two conclusions. First, there are conventional funds that invest relatively few assets in controversial stocks. And second, the fund scores are likely correlated with investment style. Another interesting finding is that we find significantly lower representation of SRI funds amongst the funds with the highest FundSTR score.

[32] Histograms 1C and 1D are harder to interpret as these scores are not between 0 and 1 due to the underlying number of ESG indicators and the adjustment procedure.

Table 3.1 Mutual fund summary statistics.

This table presents descriptive statistics on the sample of mutual funds that received scores concerning controversial investments. *FundSIN* measures the fraction of total net assets that a fund is invested in stocks associated with tobacco, alcohol, and gambling according to STATS. *FundBROADSIN* measures the fraction of total net assets that a fund is invested stocks associated with tobacco, alcohol, gambling, weapons/defense, and nuclear operations. *FundCON* (*FundSTR*) is obtained by computing for each stock *j* the fraction a fund holds in stock *j* times the total number of size and industry adjusted ESG concerns (strengths) that are reported for stock *j*, and taking the sum of the weighted scores of all stocks in the fund in a given year. The fund characteristics are presented for the period 2004-2012. A fund is defined SRI if it has at least one social investment screen. An SRI fund employs a SIN screen (SRI SIN Screen) if at least one stated screen involves tobacco, alcohol, or gambling. An SRI fund employs a BROADSIN screen (SRI BROADSIN Screen) if at least one stated screen involves tobacco, alcohol, gambling, weapons/defense, or nuclear operations. An SRI fund employs an ESG screen (SRI ESG Screen = 1) if at least one stated screen involves environmental issues. $l\_age_{i,t-1}$ is the natural logarithm of the age of the oldest share class of the mutual fund, $l\_size_{i,t-1}$ is the natural logarithm of the total net assets (TNA) of the fund, $l\_family\_size_{i,t-1}$ is the natural logarithm of accumulated TNAs of funds that belong to the same fund family, D_load_fee is a dummy for the presence of load fees, the natural logarithm of 12b1 fees, ($l\_12b1_{i,t-1}$) and that of other expenses ($l\_exp\_ratio_{i,t-1}$), past month mutual fund flow ($log\ 1+flow_{i,t-1}$) inferred from total net assets using the approach suggested in Sirri and Tufano (1998).

	Obs.	Mean	Std. Dev.	Min	Max
FundSIN	6647	0.041	0.038	0.000	0.749
FundBROADSIN	6647	0.134	0.080	0.000	1.000
FundCON	6647	-0.085	0.520	-2.372	2.308
FundSTR	6647	0.037	0.538	-2.246	3.417
Explicit SRI fund statistics					
SRI	6647	0.026	0.160	0.000	1.000
SRI SIN Screen	6647	0.021	0.144	0.000	1.000
SRI BROADSIN Screen	6647	0.022	0.146	0.000	1.000
SRI ESG Screen	6647	0.017	0.128	0.000	1.000
Fund characteristics					
Fund size	6647	1617	5809	5	161912
Family size	6647	146530	309094	5	3007970
Age (months)	6644	178	152	20	1021
Flow	6527	-0.005	0.043	-0.297	0.572
12b1	6647	0.001	0.002	0.000	0.010
Expense ratio	6647	0.010	0.006	0.000	0.089
D load fee	6647	0.521	0.500	0.000	1.000

Figure 3.1 Histograms of fund scores

Every year, we create Fund scores for all U.S. equity mutual funds in our sample with available holdings information. These fund scores measure exposures to socially sensitive business practices or industries (A=*FundSIN*, B=*FundBROADSIN*, C=*FundCON*, and *D*=*FundSTR*). In sextion 3.1 we explain how these scores are created.

Figure A: *FundSIN* histogram

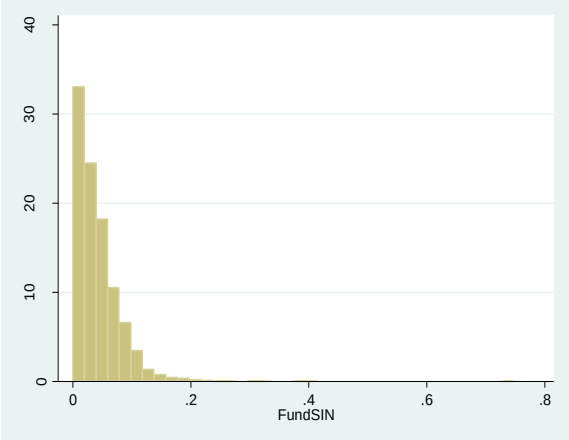


Figure B: *FundBROADSIN* histogram

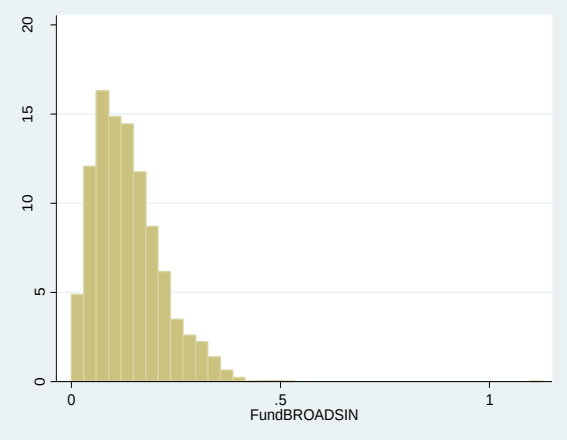


Figure C: *FundCON* histogram

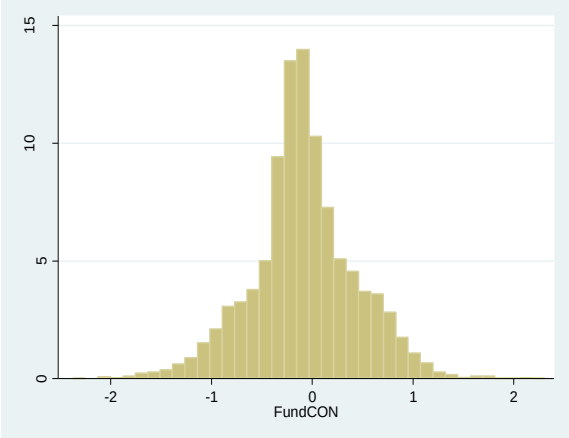
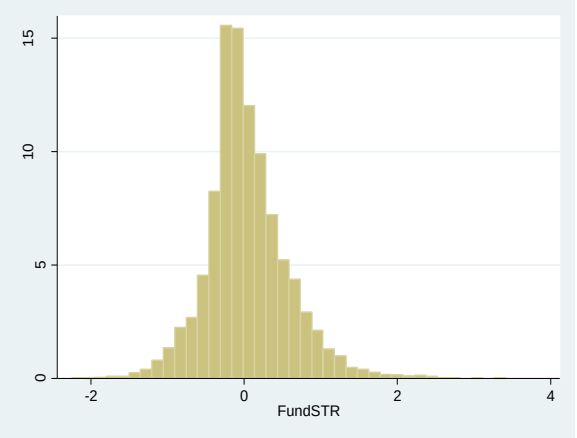


Figure D: *FundSTR* histogram



3.3.2 SRI labeled funds in the cross-section of mutual fund social exposures

Before we go into the performance effects of social tastes, we analyze the cross section of our Fund scores. Because SRI funds explicitly claim to screen their investments on the areas we consider (sin stocks, weapon and nuclear stocks, and the CSR profiles of the underlying firms as measured by ESG ratings), we expect these funds to have the lowest exposures to the sin

industries and firms with bad CSR profiles as measured by *FundSIN*, *FundBROADSIN*, and *FundCON*. Similarly we expect SRI funds to invest more in firms with strong CSR profiles as measured by our *FundSTR* score. To test this we will perform OLS regressions of the following form:

$$FundSCORE_{i,y} = c + \gamma_1 \, Dummy\ SRI_{i,y-1} + \sum_{k=2}^{K} \gamma_k \, Controls_{i,k,y-1} + \mu_{i,y} \qquad (3.3)$$

Where FundSCORE is one of our four social scores measured at the latest month available in the year for mutual fund i. Dummy SRI indicates whether the fund explicitly claims to have social investment screens. Included in the vector $Controls_{i,k,y-1}$ are fund-specific characteristics and year-fixed effects. Other controls, which are common in the literature, include a dummy variable that indicates load fees, the natural logarithm of fund age (the age of the oldest share class of the mutual fund measured in months since the inception date), fund size (the natural logarithm of total net assets (TNA) in million US$), family size (the natural logarithm of the sum of TNAs of all funds that belong to the same family), fund flow (following the approach of Sirri and Tufano (1998)), 12b1 fees, a fund's expense ratio (excluding 12b1 fees)[33], and 9 investment style dummy variables (derived by estimating sensitivities of funds' past 24-month returns to the four factors from Carhart (1997)).[34]

The results of these regressions are displayed in the first four columns of Table 3.2. We find that, in line with expectations, SRI fund portfolios have a 1.43% lower sin stock exposure on average. Considering the broader set of sin stocks, with firms in the nuclear energy and weapons industries, this difference increases to a 5.53% lower exposure.

[33] Since the fee data is on fund class level, we value-weight the fees.

[34] We create style dummies based on a 3x3 grid of investment styles, determined using the distribution of funds' Carhart (1997) factor loadings and the 33.3rd and 66.6th percentiles as cut-off levels. To estimate the factor loadings we use returns over the past 2 years and require that for each fund month we have at least 20 monthly returns. See section 3.4 for an explanation of the Carhart (1997) model.

Interestingly, the SRI funds are significantly less invested in firms with ESG concerns (t-statistic = 4.70), while these funds do not have higher exposure to firms with more ESG strengths in their portfolios. These findings support the view that SRI funds tend to avoid firms with operations that go against social tastes or norms, but do not overweight firms with strong corporate social responsibility.

All but the style controls consistently explain the exposures. The style fixed effects load significantly on the exposures as can be seen in Table 3.2. The sin stock exposure is driven by firm size exposure. Compared to the Large Blend funds, all but the Large Value funds have lower sin stock exposures and the effect is exceptionally large for Small cap funds that have 2.4% to 3.5% lower sin stock exposure on average. Similar effects can be found for the broader sin definition. On the other hand, the exposure to the adjusted ESG concerns and strengths are driven more by underlying firms' book-to-market ratios. This is not surprising as these measures are already size adjusted at the firm level. Funds classified as Growth funds (lowest 1/3rd HML beta from the Carhart (1997) fund level regressions) have lower exposures, while Value funds (Highest 1/3rd HML beta) have higher exposures compared to funds classified in the "Blend" category (Middle 1/3rd HML beta). To better control for the style effects, we run independent regressions involving model (3.3), where we allow for permutations of the abovementioned $Fund\ Score_{i,t-1}$ measures: *FundSIN*, *FundBROADSIN*, *FundCON*, and *FundSTR*, by subtracting the mean within each style group. The results, that confirm our earlier findings, are displayed in the four final columns of Table 3.2.

Table 3.2. Mutual funds socially sensitive investment exposures

We perform pooled cross-section regressions, with monthly risk-adjusted fund returns as dependent variable and as independent variables: one the funds' controversial investment measures (*FundSIN*, *FundBROADSIN*, *FundCON*, and *FundSTR*), a dummy for explicit SRI funds (SRI), a fund's age, $l\_age_{i,t-1}$ (the natural logarithm of the age of the oldest share class of the mutual fund), $l\_size_{i,t-1}$ (the natural logarithm of the total net assets (TNA) of the fund), $l\_family\_size_{i,t-1}$ (the natural logarithm of accumulated TNAs of funds that belong to the same fund family), a dummy for the presence of load fees, ($Dload\_fees$), the natural logarithm of 12b1 fees, ($l\_12b1_{i,t-1}$) and that of other expenses ($l\_exp\_ratio_{i,t-1}$), past month mutual fund flow ($log\ 1+flow_{i,t-1}$) inferred from total net assets using the approach suggested in Sirri and Tufano (1998), year-month fixed effects, and style fixed effects derived from funds' four-factor betas. *T* statistics derived from two-way clustered standard errors are presented in parentheses. Coefficients are multiplied by 100 for expositional convenience. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

	Raw fund scores				*Style adjusted fund score*			
	SIN	BROADSIN	CON	STR	SIN	BROADSIN	CON	STR
SRI	-1.425***	-5.528***	-30.346***	9.875	-1.326***	-5.295***	-27.270***	10.056
	(-2.755)	(-5.506)	(-4.700)	(1.293)	(-2.654)	(-5.696)	(-4.489)	(1.405)
Large Growth	-0.676**	-2.602***	-52.744***	-31.196***	-0.070	-0.045	-1.390	-1.754
	(-2.017)	(-7.161)	(-12.600)	(-6.628)	(-0.345)	(-0.186)	(-0.677)	(-0.679)
Large Value	0.327	1.882***	43.251***	24.108***	-0.083	0.089	-0.529	-1.188
	(1.200)	(3.710)	(8.989)	(6.505)	(-0.511)	(0.416)	(-0.270)	(-0.703)
Mid Blend	-1.160***	-2.742***	-4.389	8.055**	-0.048	0.241	0.183	-1.295
	(-4.054)	(-5.277)	(-1.038)	(2.558)	(-0.372)	(1.037)	(0.078)	(-0.719)
Mid Growth	-0.926***	-3.790***	-51.930***	-24.451***	0.039	0.300	-2.196	-0.544
	(-2.771)	(-8.021)	(-9.975)	(-3.967)	(0.224)	(0.966)	(-1.039)	(-0.250)
Mid Value	-0.996**	-2.492***	30.790***	31.706***	-0.127	0.318	-1.409	-1.320
	(-2.065)	(-2.615)	(5.970)	(5.158)	(-0.526)	(0.862)	(-0.517)	(-0.426)
Small Blend	-3.098***	-9.178***	-11.264***	-10.915***	-0.066	0.387	-2.997	-2.294
	(-6.894)	(-6.518)	(-4.484)	(-3.011)	(-0.517)	(1.264)	(-0.954)	(-0.854)
Small Growth	-2.395***	-8.775***	-22.673***	-16.286***	0.043	0.461	-4.358	-2.317
	(-5.073)	(-6.147)	(-5.816)	(-4.114)	(0.272)	(1.464)	(-1.531)	(-0.877)
Small Value	-3.521***	-9.038***	-6.256	-12.956***	0.021	0.637**	-3.917	-1.067
	(-7.872)	(-7.101)	(-1.554)	(-4.339)	(0.155)	(2.079)	(-1.041)	(-0.334)
l_age	-0.007	0.231	-2.202	-0.371	-0.003	0.241	-2.366*	-0.438
	(-0.062)	(1.285)	(-1.550)	(-0.222)	(-0.031)	(1.351)	(-1.938)	(-0.270)
l_size	-0.012	-0.211**	-0.044	-0.709	-0.018	-0.237**	0.158	-0.745
	(-0.186)	(-1.962)	(-0.060)	(-0.721)	(-0.281)	(-2.211)	(0.250)	(-0.758)
l_12b1	124.219***	164.706**	-83.944	-304.405	133.397***	201.207***	-13.771	-334.505
	(2.669)	(2.490)	(-0.168)	(-0.540)	(2.953)	(3.341)	(-0.030)	(-0.588)
l_exp_ratio	-16.512	-61.217*	-266.602	-117.859	-27.075	-102.342***	-186.265	-100.182
	(-0.681)	(-1.826)	(-1.023)	(-0.315)	(-1.130)	(-3.587)	(-0.819)	(-0.277)
D load_fees	-0.035	-0.063	2.726	-1.256	-0.054	-0.130	2.663	-0.904
	(-0.225)	(-0.270)	(1.438)	(-0.583)	(-0.359)	(-0.535)	(1.428)	(-0.405)
l_family_size	0.038	0.017	0.805*	0.471	0.028	0.003	0.705*	0.533
	(0.887)	(0.267)	(1.873)	(0.715)	(0.656)	(0.043)	(1.815)	(0.817)
Constant	5.297***	16.974***	-7.478	2.543	0.368	2.683***	-5.775	-2.431
	(5.555)	(12.633)	(-0.705)	(0.281)	(0.452)	(2.686)	(-0.598)	(-0.285)
Observations	6,287	6,287	6,287	6,287	6,287	6,287	6,287	6,287
R-squared	0.205	0.437	0.368	0.138	0.013	0.034	0.024	0.003
Time FE	YES	YES	YES	YES	YES	YES	YES	YES

Table 3.3.

See Table 3.2. We replace the SRI dummy by indicators for specific screens. An SRI fund employs a SIN screen (SRI SIN Screen) if at least one stated screen involves tobacco, alcohol, or gambling. An SRI fund employs a BROADSIN screen (SRI BROADSIN Screen) if at least one stated screen involves tobacco, alcohol, gambling, weapons/defense, or nuclear operations. An SRI fund employs an ESG screen (SRI ESG Screen = 1) if at least one stated screen involves environmental issues.

	Style adjusted fund scores							
	SIN		BROADSIN		CON		STR	
SRI SIN Screen	-1.437***		-3.217***		-2.923		6.981	
	(-3.624)		(-4.140)		(-0.562)		(0.798)	
SRI BROADSIN Screen		-1.277***		-3.314***		-2.094		9.120
		(-3.213)		(-4.134)		(-0.414)		(1.062)
SRI ESG Screen	-0.027	-0.126	-4.459***	-4.287***	-40.064***	-40.697***	10.705	8.671
	(-0.041)	(-0.179)	(-3.526)	(-3.293)	(-5.008)	(-4.969)	(1.002)	(0.812)
Observations	6,298	6,298	6,298	6,298	6,298	6,298	6,298	6,298
R-squared	0.012	0.012	0.041	0.041	0.031	0.031	0.004	0.005
Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Style FE	YES	YES	YES	YES	YES	YES	YES	YES

SRI is a multi facetted approach to investing; therefore the finding that FundSTR is not higher for SRI funds might be driven by the type of SRI screens applied. Since most SRI funds in our sample apply sin screens, some of these funds do not explicitly state to take the corporate social responsibility of firms into account. To control for the possibility that the type of screens influences our findings, we rerun the regressions with dummies for specific SRI screens.

The results, presented in Table 3.3, confirm our earlier findings and show that the screens work well since SRI (BROAD)Sin Screened portfolios have lower (broad)sin stock exposure while not affecting the exposure to ESG indicators. SRI ESG Screened portfolios have lower exposure to firms with more ESG concern indicators while not significantly affecting the exposure to sin stocks. Again, no screening method results in a higher exposure to firms with many ESG strengths.

3.3.3 Differences in mutual fund social exposures: SRI labeled vs. conventional funds

Although SRI funds might on average score lower on controversial investments, it has been suggested that the effect of social norms and values on investments are not confined to explicit SRI funds (see e.g. Morse and Shive 2011, Hong and Kostovetsky 2012, Hong and Kacperczyk 2009, Kumar, Page and Spalt 2011). In order to formally test for the possibility that conventional funds have "more social" holdings than do SRI funds, we allocate non-explicit funds to one of ten decile portfolios based on style-adjusted *FundSIN*, *FundBROADSIN*, *FundCON*, or *FundSTR*. Subsequently we compute for each decile the difference between the average style-adjusted score of all explicit SRI funds in our sample and the decile score. Table 3.4 presents these differences concerning *FundSIN*, *FundBROADSIN*, *FundCON*, or *FundSTR*, along with *t*-statistics in parentheses.

A global comparison of the first four columns (Raw fund scores) of test statistics to the second four columns (Style Adjusted fund scores) confirms our earlier findings that it is important to control for fund style when comparing exposures to controversial industries or socially desirable/controversial business practices. Controlling for fund style leads to a more fair comparison of fund scores as the differences are influenced by fund investment style.

The *t*-statistics reported in the style adjusted columns of Table 3.4 demonstrate that the group of explicit SRI funds scores significantly higher on *FundSIN* than the bottom three deciles of conventional funds, significantly higher on *FundBROADSIN* than the bottom decile of conventional funds, and significantly higher on *FundCON* than the bottom two deciles of conventional funds. For *FundSTR* we find that the top three deciles of conventional funds score significantly higher. We indicated the previously mentioned deciles by printing them in **bold**.

In addition to our SRI label, we also use dummies for the screens that match the fund scores tested (e.g. SRI Sin Screen for FundSIN). We confirm our results that there exists a

significant group of conventional funds that scores lower on *FundSIN*, *FundBROADSIN*, and *FundCON*, and there are three deciles that have a higher *FundSTR*. The magnitudes are economically significant, the top *FundSIN (FundBROADSIN)* decile of conventional funds have an 8.2% (16.7%) higher fund style adjusted (broad)sin stock exposure than SRI funds with the specific screens on average. However, the bottom *FundSIN (FundBROADSIN)* decile of conventional funds have an 2.8% (2.8%) lower fund style adjusted (broad)sin stock exposure on average.

These results demonstrate the existence of conventional U.S. equity mutual funds that are less exposed to controversial firms or more exposed to firms with ESG strengths than explicit SRI funds. This means that investors in mutual funds can consider a larger set of funds than just SRI funds when their social tastes affect their investment decisions. Especially investors into mutual funds who get utility from exposure to ESG strengths should look into the holdings of the specific funds they invest in besides following an explicit SRI label.[35]

3.4. Empirical analysis: Do social tastes affect investment performance?

3.4.1 Performance of stock portfolios potentially affected by social tastes

To understand the payoff associated with controversial stock investments witnessed in mutual fund holdings, we first perform a portfolio study on stocks that determine the fund scores. Consistent with our procedure to rank mutual funds, starting in January 2004, we build annually rebalanced stock portfolios based on previous years STATS information and collect their monthly returns. The first is a hypothetical *SIN* stock portfolio that comprises all stocks for which $Dfirm\_sin$ =1. The second is a *BROADSIN* portfolio, composed of stocks for

[35] Since most mutual funds report their top 10 holdings it is not impossible, even for individual investors into mutual funds, to assess the socially sensitive investment exposures of a mutual fund.

which $Dfirm\_broadsin$ =1. The third group of portfolios is formed on basis of the distribution of the Adj_firm_CON measure and contains all stocks in the respective quartile of the distribution. The fourth portfolio is created using a similar approach for the Adj_firm_STR measure. Doing so for consecutive years ultimately yields monthly post-rank returns from January 2004 to December 2012, which we subsequently use to determine the portfolios' average four-factor risk-adjusted return (see specification (4)).

We test the returns of these portfolios against the Carhart (1997) model. This four-factor model takes the form:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{0,i}\left(R_{m,t} - R_{f,t}\right) + \beta_{1,i}SMB_t + \beta_{2,i}HML_t + \beta_{3,i}MOM_t + \varepsilon_{i,t} \quad (3.4)$$

where $R_{i,t}$ represents each mutual portfolio's monthly return, $R_{m,t} - R_{f,t}$ is the return on a value-weighted portfolio composed of all stocks from the NYSE/AMEX/Nasdaq exchanges, in excess of a risk-free rate of return from Ibbotson. SMB_t is the return difference between a small cap portfolio and a large cap portfolio; HML_t is the return difference between a "value" portfolio (with a high book/market value ratio) and a "growth" (low book/market value) portfolio; MOM_t is the return difference between a portfolio of past 12-month winners and a portfolio of past 12-month losers.[36]

Table 3.5 shows the annualized alpha coefficient from the regressions following model (3.4) on value-weighted portfolios. Because in practice many mutual fund managers are required to hold larger firms, we consider not only hypothetical portfolios derived from stocks in the entire STATS (i.e., Russell 3000) universe but also portfolios derived from stocks that belong to the S&P500. Concerning the *SIN* portfolios reported in Panel A of Table

[36] Fama and French (1993) and Carhart (1997) provide more details on the construction of the factors and the performance evaluation model.

3.5, the average risk-adjusted return over the period 2004-2012 is 4.7% and significantly different from zero at the 5% level. A portfolio long in *SIN* stocks and short in all other Russell 3000 (S&P 500) stocks earned 5.0% (6.1%) over the same period. The value-weighted *BROADSIN* portfolio earned an economically significant risk-adjusted return of about 1.9%, which is not statistically significant at conventional levels. We report similar returns for the difference portfolios with an economically significant alpha of 2.2% for Russell 3000 and 2.3% for S&P 500 stocks.

The Adj_firm_CON and STR measures selected portfolios do not yield statistically significant returns. However, the high minus low Adj_firm_CON portfolio earned 1.5% using the Russell 3000 and 3.7% for the S&P 500 universe, both not statistically significant at conventional levels. In unreported tests we consider tercile and quintile portfolios, the results remain unchanged.

Table 3.4. Differences in scores between SRI funds and deciles of all other funds

We report the difference between the raw and style-adjusted scores of SRI funds and those of deciles of all other funds concerning the four areas of socially sensitive investment. The fund deciles are formed by ranking all non-SRI funds based on one of the three style-adjusted scores: *FundSIN*, *FundBROADSIN*, *FundCON*, or *FundSTR*. The *t* statistics on the difference in score between the SRI class and a decile is derived from a two-tailed test, and is presented in parentheses. Besides testing the difference of SRI funds and conventional funds we also test the differences using the SRI funds that specifically screen on the exposures tested. That is; *FundSIN* and SRI SIN Screen, *FundBROADSIN* and SRI BROADSIN Screen, *FundCON* and SRI ESG Screen, and *FundSTR* and SRI ESG Screen. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

	Raw fund scores				*Style adjusted fund scores*				*Style adjusted fund scores by specific screens*			
	SIN	BROADSIN	CON	STR	SIN	BROADSIN	CON	STR	SIN	BROADSIN	CON	STR
SRI-Bottom decile	0.026***	0.053***	0.675***	0.975***	0.029***	**0.037*** **	0.442***	0.899***	0.028***	**0.028*** **	**0.283*** **	0.975***
	(37.16)	(36.67)	(38.09)	(54.98)	(32.04)	**(21.60)**	(27.27)	(52.20)	(28.93)	**(16.62)**	**(15.13)**	(47.49)
SRI-2nd	0.0184***	0.025***	**0.221*** **	0.576***	0.014***	-0.000	**0.089*** **	0.499***	0.012***	-0.010***	-0.068***	0.574***
	(22.68)	(18.37)	**(16.22)**	(43.49)	(16.78)	(-0.322)	**(7.674)**	(38.94)	(14.60)	(-8.037)	(-5.741)	(40.20)
SRI-3rd	0.011***	**0.004** **	-0.004	0.438***	**0.006*** **	-0.017***	-0.063***	0.348***	**0.005*** **	-0.027***	-0.220***	0.424***
	(14.09)	**(3.046)**	(-0.324)	(34.30)	**(7.226)**	(-13.09)	(-5.638)	(27.82)	**(5.436)**	(-22.69)	(-19.98)	(30.60)
SRI-4th	**0.003*** **	-0.016***	-0.116***	0.348***	0.000	-0.030***	-0.160***	0.259***	-0.001	-0.040***	-0.317***	0.334***
	(4.174)	(-9.921)	(-9.080)	(27.58)	(0.559)	(-22.83)	(-14.54)	(20.92)	(-0.869)	(-33.93)	(-29.29)	(24.47)
SRI-5th	-0.004***	-0.037***	-0.201***	0.258***	-0.005***	-0.042***	-0.239***	0.180***	-0.006***	-0.051***	-0.396***	0.254***
	(-5.179)	(-20.03)	(-15.76)	(20.52)	(-6.425)	(-31.57)	(-21.69)	(14.64)	(-7.538)	(-44.02)	(-36.68)	(18.89)
SRI-6th	-0.013***	-0.058***	-0.285***	0.149***	-0.011***	-0.053***	-0.316***	0.093***	-0.012***	-0.062***	-0.474***	0.166***
	(-14.20)	(-28.03)	(-22.30)	(11.86)	(-13.91)	(-40.23)	(-28.72)	(7.650)	(-14.70)	(-53.85)	(-43.83)	(12.57)
SRI-7th	-0.021***	-0.080***	-0.392***	0.007	-0.018***	-0.066***	-0.404***	-0.013	-0.019***	-0.076***	-0.562***	0.061***
	(-22.95)	(-35.41)	(-30.36)	(0.502)	(-22.33)	(-49.99)	(-36.50)	(-1.037)	(-22.69)	(-64.61)	(-51.43)	(4.593)
SRI-8th	-0.031***	-0.101***	-0.562***	**-0.152*** **	-0.026***	-0.082***	-0.512***	**-0.153*** **	-0.027***	-0.092***	-0.670***	**-0.078*** **
	(-32.10)	(-43.35)	(-42.19)	**(-10.97)**	(-32.61)	(-60.93)	(-45.25)	**(-12.18)**	(-32.45)	(-76.30)	(-59.23)	**(-5.651)**
SRI-9th	-0.045***	-0.124***	-0.810***	-0.374***	-0.039***	-0.103***	-0.663***	-0.343***	-0.040***	-0.112***	-0.821***	-0.268***
	(-44.24)	(-52.71)	(-58.02)	(-24.72)	(-46.86)	(-74.58)	(-56.46)	(-25.45)	(-45.75)	(-90.31)	(-68.33)	(-17.63)
SRI-Top decile	-0.089***	-0.181***	-1.197***	-0.858***	-0.081***	-0.157***	-0.993***	-0.786***	-0.082***	-0.167***	-1.151***	-0.711***
	(-37.80)	(-53.17)	(-68.01)	(-37.67)	(-35.91)	(-55.99)	(-63.48)	(-37.53)	(-33.10)	(-56.54)	(-64.53)	(-27.87)

Table 3.5. Stock portfolios formed using controversial business indicators

Every year, starting in January we rank stocks based on controversial business indicators or ESG indicators from MSCI STATS of the previous year. Immediately following the ranking, we assign stocks to a value-weighted portfolio and rebalance the portfolio every year with the latest indicators. Doing so ultimately yields monthly post-formation returns from January 2004 to December 2012. We run Carhart (1997) four-factor regressions to estimate the risk-adjusted average return. A *SIN* portfolio contains stocks that have been associated by STATS with one of the following categories: alcohol, tobacco, and gaming. A *BROADSIN* portfolio comprises all *SIN* stocks in addition to stocks associated with military, weapons, and nuclear operations. *Adj_firm_CON* (*Adj_firm_STR*) is the size (market capitalization and market capitalization squared) and industry adjusted number of concern indicators from stats. We report on analyses performed on the entire STATS universe (Russell 3000 stocks), and on analyses performed using S&P500 firms exclusively. Panel A reports the risk-adjusted returns associated with (BROAD)*SIN*, all firms that are *Not (BROAD)SIN*, and difference portfolios (BROAD)SIN minus Not (BROAD)SIN. Panel B reports the risk-adjusted return difference of portfolios of all stocks in the respective quartiles of the distribution of *Adj_firm_CON* (*Adj_firm_STR*) as well as the difference between the Highest and Lowest quartiles. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

	Russell 3000					S&P500
Panel A: Sin stock portfolios						
	Not SIN			SIN	Difference	Difference
SIN	-0.24%			4.73%**	4.98%**	6.12%**
	(-1.123)			(2.197)	(2.164)	(2.495)
BROADSIN	-0.50%			1.85%	2.24%	2.34%
	(-1.115)			(1.202)	(1.296)	(1.266)
Panel B: Adjusted ESG score ranked portfolios						
	1/4th (L)	2/4th	3/4th	1/4th (H)	HML	HML
Adj_firm_CON	-0.53%	0.78%	-1.15%	1.01%	1.54%	3.66%
	(-0.582)	(0.768)	(-0.765)	(1.005)	(0.884)	(1.596)
Adj_firm_STR	-0.19%	0.04%	1.10%	0.02%	0.20%	-0.09%
	(-0.327)	(0.034)	(0.858)	(0.025)	(0.206)	(-0.061)

3.4.2 SR holdings and the cross-section of mutual fund performance

The analyses in the previous section lead us to expect that only sin stock exposure is associated with mutual fund returns. Although we tested hypothetical portfolios of equities,

the effects of fund scores on diversified equity funds might be different. In this section we will directly test for the effects of socially sensitive investment exposures on mutual fund performance. Our formal tests contain pooled cross-section regressions with monthly risk-adjusted fund returns $aret_{i,t}$ as dependent variable and the fund scores as the independent variables that are central to this study. Our regression models are written as:

$$aret_{i,t} = c + \gamma_1 \, Fund\ Score_{i,t-1} + \sum_{k=2}^{K} \gamma_k \, Controls_{i,k,t-1} + \eta_{i,t} \quad (3.5)$$

where $aret_{i,t}$ is the monthly return before expenses[37] of mutual fund *i* in excess of the risk free rate predicted by the Carhart (1997) factor model as in (3.4). Using this asset pricing model, $aret_{i,t}$ is then defined as:

$$aret_{i,t} = \alpha_{i,t} + \varepsilon_{i,t} \quad (3.6)$$

In independent regressions involving model (3.4), we use the abovementioned style adjusted $Fund\ Score_{i,t-1}$ measures: *FundSIN, FundBROADSIN, FundCON,* and *FundSTR.* Included in the vector $Controls_{i,k,t-1}$ are fund-specific characteristics, style fixed effects, and month-fixed effects. One particularly relevant control variable is an indicator variable for explicit SRI funds, which enables us to investigate the returns on explicit SRI funds relative to funds without SRI label. Other controls, which are common in the literature, include a dummy variable that indicates load fees, the natural logarithm of fund age (the age of the oldest share class of the mutual fund measured in months since the inception date), fund size (the natural logarithm of total net assets (TNA) in million US$), family size (the natural logarithm of the sum of TNAs of all funds that belong to the same family), fund flow

[37] We derive these returns by subtracting 1/12th of the funds annually computed expense ratio from the monthly return.

(following the approach of Sirri and Tufano (1998)), 12b1 fees, a fund's expense ratio (excluding 12b1 fees)[38]. Finally, we run rolling 24-month Carhart (1997) regressions to obtain style tilts (SMB and HML betas) and R-squared of these regressions as a measure of selectivity (Amihud and Goyenko 2013). From the betas we create 9 investment style dummy variables and include those in the analyses (for more detail see footnote 8). We derive the standard errors in two ways; first we use clusters by time and mutual fund, second we use autocorrelation-adjusted standard errors over the past 24 months following Newey and West (1987).

The results presented in Table 3.6 point to a positive association between *FundSIN* and mutual funds' risk-adjusted return[39]. The coefficient on sin-stock exposure is 0.32 (0.44 in the regressions without control variables), which suggests that an increase from zero controversy exposure to full controversial investment is associated with an annualized increase in risk-adjusted return of about 3.8% (5.3%) keeping all else equal. This return is remarkably close to the 4.7% abnormal return on the hypothetical SIN stock portfolio (Panel A of Table 3.5). The coefficient on *FundBROADSIN* is economically smaller than the coefficient on *FundSIN*, and statistically not significantly different from zero (given the conventional levels of significance). In addition, *FundCON* and *FundSTR* are not significantly related to risk-adjusted mutual fund returns over the period we investigate.

Note that in Table 3.6, the coefficient on the dummy variable for explicit SRI funds is not significantly different from zero, which is consistent with the vast majority of studies that compare the returns of SRI-labeled funds in the U.S. with those of conventional funds; see, e.g., Derwall et al. (2011) for a review of the SRI literature. The finding that risk-adjusted mutual fund return is predicted by an actual holdings-based measure of sin stock investment and not by an SRI label could be taken to imply that an SRI classification alone insufficiently

[38] Since the fee data is on fund class level, we value-weight the fees.

[39] Using raw monthly fund returns instead of risk-adjusted returns yields qualitatively similar results.

Table 3.6. Mutual funds socially sensitive investment exposures and risk-adjusted returns

We perform pooled cross-section regressions, with monthly Carhart risk-adjusted fund returns from rolling 24 month regressions as dependent variable and as independent variables: one of the funds' style adjusted socially sensitive investment measures (*FundSIN*, *FundBROADSIN*, *FundCON, and FundSTR*), a dummy for explicit SRI funds (SRI), a fund's age, $l\_age_{i,t-1}$ (the natural logarithm of the age of the oldest share class of the mutual fund), $l\_size_{i,t-1}$ (the natural logarithm of the total net assets (TNA) of the fund), $l\_family\_size_{i,t-1}$ (the natural logarithm of accumulated TNAs of funds that belong to the same fund family), the R-squared from the four-factor model over the past 24 monthly returns (R2 Carhart), a dummy for the presence of load fees, ($Dload\_fees$), the natural logarithm of 12b1 fees, ($l\_12b1_{i,t-1}$) and that of other expenses ($l\_exp\_ratio_{i,t-1}$), past month mutual fund flow ($l\_flow_{i,t-1}$) inferred from total net assets using the approach suggested in Sirri and Tufano (1998), year-month fixed effects, and style fixed effects derived from funds' rolling 24 month four-factor betas. *T* statistics derived from two-way clustered standard errors are presented in *round* brackets and derived from Newey-West corrected standard errors with 24 lags in *squared* brackets. Coefficients are multiplied by 100 for expositional convenience. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

	Dependent variable risk adjusted return							
Fund score:	SIN		BROADSIN		CON		STR	
Fund score	0.444	0.324	-0.108	-0.140	0.002	-0.009	-0.003	-0.006
	(1.826)*	(1.316)	(-0.606)	(-0.781)	(0.042)	(-0.195)	(-0.120)	(-0.227)
	[2.391]**	[1.708]*	[-0.911]	[-1.157]	[0.142]	[-0.661]	[-0.277]	[-0.529]
l_flow		0.510		0.520		0.516		0.517
		(1.860)*		(1.898)*		(1.887)*		(1.886)*
		[2.783]***		[2.835]***		[2.815]***		[2.820]***
R2 Carhart		-0.706		-0.735		-0.735		-0.731
		(-1.543)		(-1.617)		(-1.654)*		(-1.614)
		[-4.520]***		[-4.754]***		[-4.762]***		[-4.716]***
SRI		-0.007		-0.018		-0.013		-0.010
		(-0.124)		(-0.330)		(-0.258)		(-0.184)
		[-0.179]		[-0.472]		[-0.352]		[-0.266]
l_age		-0.005		-0.004		-0.005		-0.005
		(-0.509)		(-0.455)		(-0.536)		(-0.493)
		[-0.552]		[-0.497]		[-0.554]		[-0.532]
l_size		-0.009		-0.009		-0.009		-0.009
		(-1.865)*		(-1.964)**		(-1.884)*		(-1.872)*
		[-1.906]*		[-1.974]**		[-1.920]*		[-1.927]*
l_12b1		-7.590		-6.909		-7.168		-7.173
		(-1.845)*		(-1.686)*		(-1.736)*		(-1.738)*
		[-1.971]**		[-1.795]*		[-1.867]*		[-1.868]*
l_exp_ratio		-1.646		-1.975		-1.809		-1.819
		(-0.289)		(-0.350)		(-0.318)		(-0.320)
		[-0.379]		[-0.455]		[-0.418]		[-0.420]
Dload_fees		0.021		0.021		0.021		0.021
		(1.166)		(1.128)		(1.165)		(1.139)
		[1.376]		[1.333]		[1.363]		[1.344]
l_family_size		0.012		0.012		0.013		0.012
		(2.374)**		(2.403)**		(2.465)**		(2.415)**
		[4.134]***		[4.171]***		[4.199]***		[4.180]***
Constant	0.017	0.706	0.017	0.735	0.017	0.734	0.017	0.730
	(0.463)	(1.822)*	(0.464)	(1.912)*	(0.463)	(1.966)**	(0.464)	(1.907)*
	[2.930]***	[4.140]***	[2.933]***	[4.341]***	[2.932]***	[4.344]***	[2.932]***	[4.304]***
Observations	52,321	51,984	52,321	51,984	52,321	51,984	52,321	51,984
R-squared	0.000	0.081	0.000	0.081	0.000	0.081	0.000	0.081
Time FE	NO	YES	NO	YES	NO	YES	NO	YES
Style FE	NO	YES	NO	YES	NO	YES	NO	YES

distinguishes mutual funds along social dimensions.[40] Section 4 will more explicitly elaborate on the usefulness of the SRI label relative to that of holdings-based information in assessing the effects of social preferences or tastes on risk-adjusted return.

Given the precise definition of what stocks are considered sin stocks in our paper, the *FundCON* and *FundSTR* measures are derived indirectly from firm level ESG indicators. Therefore, we violate the assumption that the independent variables are measured without error. To mitigate this issue we will consider several measures that suffer less from estimation error. The first measure considered is when we use quintiles of the *firm-level* distributions of Adj_firm_CON and STR to form *FundCON* and *FundSTR* as in (1c) and (1d). The results presented in Table 3.7 confirm our findings from Table 3.6 that there is no significant effect of ESG concerns on risk-adjusted returns of diversified U.S. equity funds. However, *FundSTRgroups* is significantly negatively related to mutual fund performance in 3 out of 4 regressions. This result is in line with tastes for stocks of companies with strong corporate social responsibility driving up the prices of these companies and consequently leading to lower expected returns.

Second we test for the possibility that social tastes for firms with much ESG strength indicators are particularly positive when these firms do not have a lot of concerns. A similar argument can be made for social tastes against firms with a high number of concerns, those firms might try (and succeed) to compensate with a high number of ESG strengths. Therefore we create firm level dummies that are equal to one when a firm is in the top (bottom) 40% of Adj_firm_CON and in the bottom (top) 40% of Adj_firm_STR to form the fund-level measures: FundCONtop_STRbottom (FundCONbottom_STRtop). We repeat this exercise for top and bottom 20% firms. The results indicate that the negative returns associated with exposure to firms with more strengths are driven by firms in the top quintile of strengths and

[40] The coefficient on the SRI dummy continues to be not significantly different from zero if we drop the three holdings-based measures of controversial investment.

bottom quintile of concerns. Broadening this definition to the top and bottom two quintiles results in all coefficients losing their statistical significance and decreasing in magnitude as well.

Table 3.7 Mutual funds ESG rating exposure and risk-adjusted returns

See Table 3.6. We consider variations on the *Fund scores*. First we replace Adj_firm_CON (Adj_firm_STR) with quintiles of Adj_firm_CON (Adj_firm_STR), firms in the third quintile receive a firm level score of 3. We use these quintile measures to rank the mutual funds as in (1c) and (1d) to form FundCONgroups and FundSTRgroups. Second we consider firm level dummies for firms that are both in the Top (Bottom) quintiles of Adj_firm_CON *and* in the bottom (Top) quintiles of Adj_firm_STR to form Fund *CONtop STRbottom* (*FundCONbottom STRtop*). We perform this exersize for firms in the Top and Bottom 40% as well as for the Top and Bottom 20%.

	Dependent variable risk adjusted return					
FundCONgroups	-0.006	-0.013				
	(-0.182)	(-0.388)				
	[-0.473]	[-1.001]				
FundSTRgroups	-0.040	-0.037				
	(-1.701)*	(-1.604)				
	[-2.516]**	[-2.379]**				
Fund CONtop STRbottom			0.037	0.005		
40%			(0.154)	(0.020)		
			[0.274]	[0.036]		
FundCONbottom STRtop			-0.174	-0.145		
40%			(-0.902)	(-0.768)		
			[-1.241]	[-1.082]		
FundCONtop STRbottom					-0.347	-0.383
20%					(-0.671)	(-0.750)
					[-1.215]	[-1.366]
FundCONbottom STRtop					-0.625	-0.475
20%					(-1.789)*	(-1.492)
					[-2.402]**	[-1.903]*
Observations	52,321	51,984	52,321	51,984	52,321	51,984
R-squared	0.000	0.081	0.000	0.081	0.000	0.081
Month FE	NO	YES	NO	YES	NO	YES
Style FE	NO	YES	NO	YES	NO	YES
Controls	NO	YES	NO	YES	NO	YES

Taken together, this section provides evidence that social tastes can affect returns. The findings in this section are consistent with the notion of Hong and Kacperczyk (2009) that in particular investor boycotts of classical sin stocks from the tobacco, alcohol and gaming

sectors are material enough to influence asset prices. Similarly, our findings are consistent with tastes for socially responsible firms leading to lower returns for firms with many strength indicators and few concern indicators. However, although the estimated *payoffs* that mutual funds can enjoy for a full unit of investment exposure is large, it is not entirely clear at this point to what extent mutual funds realized higher (lower) returns by taking full *exposure* to controversial (socially responsible) assets. How much factor exposure mutual funds adopted throughout our sample period and the effects on realized pre-expense returns is addressed in the next section of the paper.

3.4.3 Socially sensitive investment and performance of portfolios of mutual funds

We now investigate to what extent mutual funds' risk-adjusted returns are affected by the degree of their actual investment in socially sensitive securities. Every year, we rank all funds on one of their style-adjusted fund scores and then allocate funds to several ranked mutual fund portfolios. In independent analyses, we rank funds on one of the four style-adjusted scores of socially sensitive investment. We collect monthly returns on each portfolio for the next twelve months. By annually ranking funds using updated mutual fund scores, we eventually obtain monthly returns on mutually exclusive portfolios, which differ along the aforementioned scores, for the period January 2004-December 2012. Our principal goal is to assess the risk-adjusted return difference between high- and low-ranked mutual funds, where the risk-adjusted average return is based on the intercept from a four-factor model (as in specification (3.4)). If mutual funds can profit from social controversies, we expect to see a positive risk-adjusted return difference between high-ranked and low-ranked ("high minus low") portfolios. Similarly, funds with the highest exposures to firms with good CSR profiles might gain lower risk adjusted returns than their low exposure peers.

Table 3.8 shows the results of ranking funds based on the style-adjusted fund scores and assigning the highest (lowest) one-fourth of all ranked funds to high (low) ranked portfolios.[41] In Panel A of Table 3.8, we show the results of portfolios where mutual funds receive a weight in the portfolio based on their total net assets (TNA) relative to the TNA of all funds in the portfolio in addition we provide high minus low (HML) equal weighted portfolios of mutual funds.

The results in Panel A of Table 3.8 show that the annualized risk-adjusted return differences between top- and bottom-ranked funds are small and not statistically significant for all fund-scores we consider. The observation that the risk-adjusted return differences between top- and bottom-ranked funds in Table 3.8 are smaller than the payoff per unit of fund score exposure as reported in Tables 6 and 7, is arguably because mutual funds with relatively high scores are still not fully invested in sin stocks or stocks with many ESG strengths due to benchmark (tracking error) constraints and other trading restrictions as shown in Table 3.4. Evidently, the spread in mutual fund scores between top and bottom-ranked funds is considerably smaller than the 100-percent spread that is implicitly required to reap the estimated payoffs that are reported in Tables 6 and 7.

In Panel B we complete this analysis by a consideration of the exposures to firms with good (poor) social responsibility profiles, defined as having many (few) ESG strengths and few (many) concerns. In line with the cross-sectional analyses we find that a TNA weighted (equal weighted) High minus Low portfolio of funds ranked on their exposure to firms in the top 40% of Firm_adj_STR *and simultaneously* in the bottom 40% of Firm_adj_CON yields a negative alpha of -0.90% (-0.55%) which is not statistically significant. When considering the very best corporate citizens by considering top and bottom 20% cut offs, the statistical significance increases and we find a negative alpha of -1.08% on the TNA weighted portfolio

[41] We also considered tercile and quintile portfolios of mutual funds, the results remain virtually unchanged for these fund selection procedures.

that is significant at the 10% level. The equal weighted counterpart is not statistically significant at the levels we consider.

In a nutshell, the exposure differences between mutual funds are not large enough to lead to convincing significant performance differences between high vs. low exposure groups of funds. Only sporadically do we find funds that are fully invested in controversial or most socially responsible assets. Social tastes have the potential to influence investment performance, however, the exposures are too low for us to find large and consistent differences between large numbers of mutual funds.

3.4.4 Socially sensitive investment exposures and performance of SRI labeled funds

The analysis of mutual fund holdings indicates that the extent of investment in sin stocks and socially progressive stocks influences mutual fund performance. At first glance, this conclusion appears to contrast with studies on the performance of explicitly socially responsible mutual funds, also known as SRI funds. The typical U.S. fund with SRI label explicitly conducts exclusionary screens on stocks from tobacco, alcohol and gambling sectors, among other types of social investment screens. Yet the majority of empirical studies on SRI-labeled mutual fund performance concludes that SRI funds and conventional funds earn similar risk-adjusted returns; see, e.g., Derwall et al. (2011) for a review of empirical studies. In Table 3.9, we corroborate this conclusion based on our sample of SRI funds over the period 2004-2012. Reported is the four-factor alpha for a portfolio of all SRI funds relative to that of a portfolio of funds that do not carry an SRI label, as well as the alphas for portfolios of SRI funds that employ specific socially responsible investment screens. The alphas in Table 3.9 are not significantly different from zero. Thus, unlike evidence from actual holding in socially sensitive assets, evidence from SRI-labeled funds and their stated

investment screens would suggest that responsible investment screens do not influence fund performance.

Table 3.8. Risk-adjusted returns of exposure ranked fund portfolios: alternative cut-off levels of socially sensitive investments

Every year, we rank all mutual funds in our sample for which holdings information is available on their most recent Fund score, using one of four style-adjusted measures (*FundSIN*, *FundBROADSIN*, *FundCON*, and *FundSTR*). We style adjust the scores by subtracting the mean of the score within each style group. Immediately following the ranking, we assign funds with high (low) scores to a portfolio composed of Top (Bottom) ranked funds. We form quartile portfolios based on the Fund score distributions of all ranked funds. We compute the portfolios' monthly returns for the next twelve consecutive months. This procedure ultimately yields monthly post-formation returns from January 2004 to December 2012. We run Carhart (1997) four-factor regressions to estimate the risk-adjusted average return. We report the complete set of results for value weighted portfolios based on total net assets (TNA) and the difference portfolios (HML) for both TNA and equal weighted scemes. *T* statistics are presented in parentheses.

	TNA weighted					Equal weighted
	1/4th (L)	2/4th	3/4th	1/4th (H)	HML	HML
Panel A: Style adjusted scores						
SIN	0.34%	-0.78%	-0.63%	0.49%	0.15%	-0.06%
	(0.669)	(-1.480)	(-1.217)	(0.757)	(0.334)	(-0.169)
BROADSIN	0.49%	-0.07%	-0.11%	-0.26%	-0.75%	-0.08%
	(0.882)	(-0.0966)	(-0.204)	(-0.415)	(-1.267)	(-0.215)
CON	0.19%	-0.25%	-0.49%	0.39%	0.20%	0.18%
	(0.293)	(-0.355)	(-0.864)	(0.613)	(0.275)	(0.302)
STR	0.22%	0.35%	-1.09%	0.23%	0.01%	-0.40%
	(0.358)	(0.586)	(-1.938)*	(0.361)	(0.0165)	(-0.859)
Panel B: Stlye adjusted Top vs. Bottom STR and CON scores						
CONtop STRbottom 40%	-0.14%	0.22%	-0.81%	0.73%	0.87%	0.22%
	(-0.201)	(0.328)	(-1.454)	(1.322)	(1.069)	(0.469)
CONbottom STRtop 40%	0.69%	-0.33%	-0.42%	-0.21%	-0.90%	-0.55%
	(1.088)	(-0.547)	(-0.695)	(-0.353)	(-1.593)	(-1.330)
CONtop STRbottom 20%	-0.42%	0.20%	-0.35%	-0.04%	0.38%	-0.25%
	(-0.781)	(0.297)	(-0.599)	(-0.0674)	(0.606)	(-0.720)
CONbottom STRtop 20%	0.94%	-0.39%	-0.38%	-0.14%	-1.08%	-0.49%
	(1.502)	(-0.650)	(-0.632)	(-0.234)	(-1.771)*	(-1.433)

Table 3.9. Difference in four-factor alphas between SRI-labeled funds and all other funds

Reported are four-factor risk-adjusted returns based on specification (3.3), derived from monthly returns of explicit SRI fund and all other ("no-SRI") funds in the sample over the period January 2004 – December 2012. Panel A reports on a total net assets (TNA) weighted portfolio of SRI-labeled funds minus a portfolio of all other funds in the sample, and on an equal-weighted portfolio of SRI funds that employ sin screens minus a portfolio of all funds that do not employee sin screens. Panel B reports four-factor alphas concerning these fund portfolios when funds receive equal portfolio weights. *T* statistics are presented in parentheses. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

	Alpha	Mkt-Rf	SMB	HML	MOM	R-squared
Panel A: TNA weighted						
SRI - No SRI	-0.77%	-0.053**	0.066*	0.098***	-0.053***	0.228
	(-0.833)	(-2.310)	(1.686)	(3.355)	(-3.454)	
Sin Screen – No Sin Screen	-0.77%	-0.052**	0.063	0.098***	-0.053***	0.224
	(-0.832)	(-2.274)	(1.612)	(3.359)	(-3.380)	
ESG Screen – No ESG Screen	-1.30%	-0.028	0.110**	0.104***	-0.085***	0.250
	(-1.165)	(-0.836)	(2.274)	(2.974)	(-3.510)	
Panel B: Equal weighted						
SRI- No SRI	-0.53%	-0.049**	-0.036	0.056*	-0.034***	0.200
	(-0.748)	(-2.108)	(-1.103)	(1.816)	(-3.078)	
Sin Screen- No Sin Screen	-0.70%	-0.037*	-0.049	0.058**	-0.029***	0.207
	(-1.067)	(-1.835)	(-1.656)	(2.138)	(-2.863)	
ESG Screen - No ESG Screen	-1.08%	-0.031	-0.004	0.050*	-0.069***	0.338
	(-1.296)	(-1.282)	(-0.091)	(1.664)	(-3.591)	

However, one explanation that reconciles these different findings is that various non-explicit SRI funds experience lower controversial scores than the typical SRI fund (Table 3.4). Although SRI funds might on average score lower on controversial investments, it has been suggested that the effect of social norms and values on investments are not confined to explicit SRI funds (see e.g. Morse and Shive 2011, Hong and Kostovetsky 2012, Hong and Kacperczyk 2009, Kumar et al. 2011). According to our findings, excluding sin stocks and overweighting socially progressive stocks (many strengths and few concerns) both could

result in lower SRI fund performance. However, in Tables 2 to 4 we report that the exposure differences between SRI and conventional funds are generally small mitigating the effects of social tastes on SRI fund performance.

These findings have implications for determining whether investors can profit from social tastes in markets with the use of mutual funds. While the predominant approach in the literature involves contrasting explicit SRI funds with conventional funds, our results imply that such a comparison leads to biased inferences about the practical implications of social dimensions regarding investment performance. If several conventional funds score at least as low as SRI funds in terms of controversial investments, then a comparison of returns between SRI and conventional funds will mask the true relation between social dimensions in fund holdings and fund performance. These results reinforce the case for studying actual fund holdings to determine the true social responsibility profiles of mutual funds' investments.

3.5. Conclusion

Theories predict that due to social tastes a significant number of investors abstain from holding stocks that are deemed socially controversial or overweight stocks from socially responsible firms (e.g. Fama and French 2007, Hong and Kacperczyk 2009). So far, there has been mixed empirical evidence on the question whether these social tastes have an impact on the prices and returns of socially sensitive stocks. Can investors profit from social tastes?

We started our analyses with a comparison of fund exposures to socially sensitive stocks. We show that there is significant variation in the extent to which funds are exposed to Alcohol, Tobacco, and Gambling firms ("sin" firms), a broader sin definition that also includes Weapons, and Nuclear energy stocks, as well as variation in exposure to firms with a high number of ESG concern or strength indicators. As expected, SRI funds hold fewer assets

in controversial industries and firms with ESG concerns on average. However, an investigation of the distribution reveals that there exists a large group, ranging from 10% to 30%, of the conventional mutual funds that is less exposed to these types of "bad" firms. On the contrary, they do not differentiate themselves on the level of social and environmental strengths.

These findings support our view of a holdings-based analysis instead of comparing SRI labeled funds to conventional funds to test the effects of social tastes on investment performance. Based on the actual holdings and pre-expense returns of domestic U.S. equity mutual funds over the period 2004-2012, we provide evidence that social tastes can affect investment performance, yet to a very limited extent. This finding contrasts studies based on hypothetical stock portfolios with full exposure to socially sensitive stocks (like sin stocks) that have provided economically large and statistically significant returns (e.g. Fabozzi et al. (2008), Hong and Kacperczyk (2009), and Derwall et al. (2011)).

We find that differences in mutual funds' risk-adjusted returns can be explained by differences in their 'sin stock' scores (the fraction of assets invested stocks from alcohol, tobacco, and gambling sectors), and their exposure to socially responsible firms (those that simultaneously score high on our measure of ESG strength indicators and low on the concerns measure). We conclude that the effect of social tastes on asset prices is economically and statistically significant in the presence of actual trading costs and trading restrictions for sin stocks, as predicted by Hong and Kacperczyk (2009), and for stocks with very strong social responsibility profiles as predicted by El Ghoul et al. (2011). Beyond these social dimensions, we do not find evidence that exposure to a broader set of sin stocks and stocks with ESG concerns have influenced risk-adjusted mutual fund returns.

To build the bridge from these cross-sectional results to a profitable investment strategy we should find an abnormal risk adjusted return for a portfolio long (short) in mutual funds

with high (low) sin stock exposure. Or a portfolio of mutual funds that is short (long) in firms with a high (low) number of ESG strengths and a low (high) number of concerns. Our evidence however shows that the cross-sectional results do not work through into high minus low portfolio returns due to low exposure differences between funds. Subsequently, we use a similar logic to explain why the literature has found that SRI fund performance is not different from conventional fund performance even though they explicitly state to exclude (profitable) controversial stock investments.

Although the performance differences between SRI and conventional funds are small, we do find some potentially interesting avenues for further research into SRI funds. The combined findings that SRI funds do not overweight firms with a higher number ESG indicators, and that it is exactly this exposure that leads to (slightly) lower risk adjusted returns, could support the conclusion that SRI funds better understand the effects of corporate social responsibility on performance and therefore do not overweight firms with very strong corporate social responsibility policies. On the other hand, SRI funds can also act as merely social concern avoiders like they do, to some extent, for sin stocks.

Taken as a whole, a holdings-based analysis of mutual funds provides valuable new insights into the factor exposures and payoffs associated with socially sensitive assets in practice. Based on our main findings we suggest using actual holdings data for analyses into the effects of tastes on asset prices and returns.

Chapter 4

4. Attitudes towards socially and environmentally responsible investment[42]

4.1. Introduction

There has been an increasing interest in the effects of social values on investment behavior and decision making. In this paper we empirically test whether the utility obtained from investing in socially responsible pension investments is significantly positive. If so, pension funds face the challenge of implementing a responsible investment strategy on behalf of their beneficiaries. If this is the case, we first need to investigate if and to what extent individuals value social responsibility in their pension investments because beneficiaries do not have full information on and control over their pension investments. At the same time the pension funds have significant market power.

The rationale behind the introduction of more socially responsible pension products initiates in the observation that people are increasingly buying socially responsible products. The market share of green energy in the Netherlands has risen from below 2% in 1990 to over 11% in 2008[43]. The UK Fair-trade foundation[44] claims to have labeled over 3000 products as Fair trade in 2011. From 2009 onwards more than 100 million pounds of Fair Trade certified coffee was imported into the U.S. alone of which 62% was organic[45]. These numbers reveal non-ignorable preferences towards socially responsible consumer product features.

[42] This Chapter also circulates as Borgers and Pownall (2014).

[43] http://www.eia.gov

[44] http://www.fairtrade.org.uk

[45] http://www.transfairusa.org/sites/default/files/Almanac%202010_0.pdf

Nevertheless, the current economic literature on pensions focuses on risk and return. Utility from non-pecuniary aspects is largely left out of consideration while the pension fund managers should maximize the utility of their beneficiaries. If individuals on average derive positive utility from more socially responsible pension investments, pension fund managers should incorporate these values of their beneficiaries into their investment schemes. This is not an easy task, Barber (2007) points out: "Once considerations other than wealth maximization are relevant for investors, aligning the interests of portfolio managers and investors becomes extremely difficult". Think about regulatory difficulties concerning heterogeneous preferences (Richardson 2011) or the interpretation of fiduciary duties (Richardson 2007, 2009). However, these difficulties cannot be an argument to refrain from taking environmental and social preferences into account at all.

Despite theoretical difficulties, we have seen in practice that tobacco stocks are excluded from investment portfolios in the U.S. (e.g. CalPERS, LA, Seattle, Chicago, and NY pension fund), some of these funds also exclude assault weapon related companies, or companies with business in Sudan. A number of Northern European pension funds have committed to responsible investment practices like the exclusion of cluster bomb manufacturers. The two largest pension funds in the Netherlands (ABP and PGGM) divested Walmart in 2013 due to unsuccessful engagement with the firm concerning questionable employee relations. These examples show that pension funds do take into account social responsibility, although, these practices are also highly debated. Restrictions on the investable universe can harm risk adjusted returns which is not in line with the fiduciary duty of pension

funds and the beneficiaries within funds have different social preferences which can lead to protests[46].

The examples above show that some pension funds take social preferences into account in pension fund investments. However, in the conventional pension fund setting, participants do not invest themselves. Pension funds fulfill a fiduciary role in that they invest on behalf of their participants. Because the participants have very limited power to influence the investment decisions of fund managers and the mandates dictate fund managers to invest prudently (meaning highest possible return for lowest possible risk) there is an inherent agency problem whenever a socially desired investment choice would have a negative impact on expected returns. For this reason it is important to understand if beneficiaries want their money to be invested more in line with their social preferences. The first aim of this paper is to provide descriptive information on attitudes towards social responsibility in pension investments. Subsequently, we compare stated preferences by investigating willingness to pay for socially responsible investments while taking these attitudes into account.

Taking preferences into account makes the investment strategies of fund managers extremely difficult due to heterogeneity (Barber 2007). A straightforward solution would be to give beneficiaries greater freedom of choice or even full investor autonomy. This is only possible if investors are financially responsible and capable of making sound financial decisions. Our second aim is to understand to what degree beneficiaries are able to translate their stated preferences into financial decisions.

Using unique field data on Dutch households from CentERdata that was gathered in the first quarter of 2011 we are the first to investigate environmental and social preferences of

[46] E.g. when two Dutch pension funds (PGGM and ABP) donated 1.5 million Euro to IFKO (an international fund for vulnerable elderly) after the tsunami of 2004, complaints were made by their beneficiaries in different types of media. They did not want the pension fund to decide *how* their money should be donated.

pension fund beneficiaries. We investigate the extent to which beneficiaries claim to value several environmental and social characteristics of companies and test if they are able to translate these preferences into financial decisions consistently. Subsequently we validate our findings by looking into the effects of social and environmental screening on utility. Our contributions can be separated into three categories.

First, we compare attitudes of Dutch household members in 2011 towards socially responsible investment practices. In contrary to the conventional practice in the U.S. it would not be optimal to exclude the sin companies (alcohol, tobacco, and gambling) from pension investments. On average, the Dutch favor the exclusion of companies that violate human rights and companies operating in the weapons industry. Although there are some commonalities, there is significant variation in the attitudes of the beneficiaries. To validate the reported scores we show that the respondents were able to express their attitudes towards social responsibility screens as their self assessed levels correspond with their daily behavior, such as those who consume alcohol or smoke.

Second, we report that the respondents experience difficulties making financial decisions while at the same time taking their attitudes towards social responsibility screens into account. Specifically, over one third of the respondents reported decisions inconsistent with rational behavior. We show that these inconsistencies are partially explained by low levels of financial sophistication.

Third, to measure the effect of socially responsible in- vestment screening on utility we examine willingness to pay and find that around three quarters of the respondents are willing to give up pension income to get their investments more in line with their social and environmental preferences. The likelihood to be willing to pay is lower for men, and rises in education, income, and especially in having a positive attitude for social and environmental screening, the latter leads to a 40% increase in willingness to sacrifice pension income. These

stated preferences are not easily explained, however, results lead us to advise pension funds to reveal social and environmental preferences by directly engaging with beneficiaries.

Our research suggests a number of important points for policymakers and pension funds which should optimally execute their fiduciary duty. Finally, the paper can help ethical investment funds to identify potential customers.

The organization of the paper is as follows. In Section 2, we show that the literature finds that people do not only wish to maximize wealth but are likely to value social aspects in their pension investments. We question if respondents might even want to give up income for the socially responsible dimension in their pension, though, individuals are likely to have heterogeneous preferences. The final part of Section 2 explains that the low financial literacy of household members imposes restrictions on the degree of pension portfolio customization. Section 3 presents and describes the data used. Section 4 presents tests on non-financial preferences and financial decision- making. The fifth Section validates the paper by showing that beneficiaries on average have utility defined over the level of social responsibility in their pension investments. Section 6 provides a discussion and Section 7 concludes.

4.2. Theory

4.2.1. Willingness to pay for social and environmental screening

The literature on pension fund investments is focused on risk and returns while non-pecuniary utility is largely left out of consideration. This is unsurprising as pension funds exist foremost to provide income after retirement for their beneficiaries. However, most individual investors consider investments as consumption goods and not just as investment products (Keloharju et al. 2012), meaning that non-financial attributes can influence the decision to invest or not.

This view towards investments can explain findings of socially responsible investment research. Segments of mutual fund investors gain non-pecuniary benefits from investing in socially responsible investments (e.g. Bollen 2007, Nilsson 2009), the same holds for socially responsible banking clienteles (Bauer and Smeets 2013). From owner- ship studies we know that such investors are less focused on past return performance (e.g. Renneboog et al. 2011). Analogously, the consumer literature finds segments of consumers with varying attitudes to or stated preferences for ethical product features. For example, Auger et al. (2008) report that within their specific sample 40% is classified as a socially conscious buyer of soap bars and/or athletic shoes. Furthermore, De Pelsmacker et al. (2005) report that within their sample 50% of consumers attached importance to the fair trade attribute of coffee. These findings together with the evidence on the socially responsible investor and banking clienteles lead to our first hypothesis.

H1: Beneficiaries have heterogeneous attitudes towards socially responsible pension investments.

To justify social responsibility in pension investments, it is important to investigate whether the attitudes of beneficiaries factor into their utility functions. This can be done by analyzing willingness to pay for changing pension investments to be more in line with the reported attitudes. For social choices to be reflected into investments, fund managers need to collect information on firms' activities and in addition try to reveal preferences of the individuals and take them into account in the best possible way. Given the preferences of the beneficiaries, the manager can either screen companies and/or engage with companies to change their corporate behavior. For simplicity we will focus on screening throughout the paper.

Screening is selecting companies to invest in with certain criteria, for instance based on social or environmental performance measures. The investment literature reports mixed results for the effects of screening on returns. In theory the screening process itself is costly, in addition limiting the investment space might harm returns (Hong and Kacperczyk 2009, Statman and Glushkov 2009).[47] Keeping all else equal, returns drop as the investable universe decreases due to the responsible investment constraint (Boudt et al. 2013). However, empirical evidence also shows companies that engage in different forms of social responsibility beat their benchmarks in the stock market but only for specific screens (e.g. Derwall et al. 2005, Kempf and Osthoff 2007). Borgers et al. (2013) show that these positive effects vanish in the long term as attention for social and environmental business practices increases. Consequently, the negative effects may prevail in the long term. However, we acknowledge that even if the screening process is costly, the effects for individual pension beneficiaries are negligible as long as the screens do not significantly reduce the investable universe. Therefore, we want to stress here that we merely use willingness to pay as a means to test if the common mans' utility is at all influenced by social and environmental factors in investments.

Empirical studies that have looked into willingness to pay report that significant segments of consumers are readily willing to pay for non-product environmental (Laroche et al. 2001) or ethical (Auger et al. 2003, 2008) features. In an experimental study Andreoni and Miller (2002) identify a quarter of the investigated population as pure money-maximizers which implies that three quarters do give up income for non-pecuniary utility. Since the

[47] We assume that pension fund managers (try to) maximize returns within the freedom of their mandates. The mandates make sure that the investments comply with prespecified demands of governments and other stakeholders. This means that if the pension investments are further screened on social issues the investment performance will at best be the same as in the current situation since screening will limit the investment space. The same reasoning can be applied for engagement strategies. If engagement adds to the risk return performance of the pension investments, the manager should already be doing this.

amount of pension investments is enormous, it is relevant to investigate if pension beneficiaries gain utility from incorporating their non-financial preferences into their pension investments. One way to do this is by looking at willingness to pay for socially responsible investment screens. Leading to our second hypothesis.

H2: Beneficiaries are willing to give up pension income to better align their pension investments with their social attitudes.

4.2.2. Financial literacy and socially responsible investing

If beneficiaries derive positive utility from more socially responsible pension investments and the preferences are heterogeneous, a straightforward solution is to increase the level of investment autonomy. Allowing the common man to have greater freedom of choice on his preference for responsibly investing pension fund money on his behalf. This strongly requires a high level of financial literacy. Therefore, we investigate if beneficiaries are actually able to make financial decisions while at the same time taking their non-financial preferences into account. Currently the Dutch pension system works with defined benefits with very limited influence from beneficiaries.

Moreover, Benartzi and Thaler (2005) find that many defined contribution pension funds doubt the quality of investment strategies of their beneficiaries. Van Rooij et al. (2007) show that people are on average not able to make consistent choices in a risk-return tradeoff universe. This effect is even stronger for beneficiaries with limited financial knowledge.

In this respect, it is important to take financial literacy into account as more financially literate individuals are more involved with their financial decisions and make more sophisticated financial choices. Some examples, they are more prone to plan for their retirement (Van Els et al. 2004, Lusardi and Mitchell 2007a,b, Van Rooij et al. 2011a), they

hold more diversified portfolios (Calvet et al. 2009a,b), have higher levels of savings (Bernheim et al. 2001, Bernheim and Garrett 2003), and they are more likely to participate in the stock market (Van Rooij et al. 2011b).

As we use a sample of Dutch households, we know from Van Rooij et al. (2011a,b) that the respondents do under- stand basic financial and economic concepts, though, when financial products become more complex there is very little knowledge at hand. Since, introducing a social responsibility attribute to the choice framework complicates financial decisions by introducing an extra dimension we expect that the average respondent will not be able to make financial choices while simultaneously taking their reported attitudes into account. A consequence would be that giving respondents more freedom of choice is not the optimal way to implement their non-financial preferences into their pension investments. This reasoning is in line with our third and fourth hypotheses.

H3: Beneficiaries are unable to make consistent financial decisions while simultaneously taking their non-financial preferences into account.

H4: Beneficiaries with low levels of financial literacy are less capable of making financial decisions while simultaneously taking their non-financial preferences into account.

4.3. Data

To test our hypotheses we use a unique dataset on Dutch households. The data are derived from a customized questionnaire, matched to a wide range of demographics from the CentERdata Databank at Tilburg University. Respondents are members of the CentERpanel[48]

[48] Our survey is sent out once so we cannot benefit from the panel structure of the data. We translate the questions used throughout this paper into English and present them in appendix A.

who participate in weekly surveys over the internet using a computer. When a computer is not available, the members are provided with a television set up box which makes the sample selection exogenous from the availability of an internet connection. The members of the CentERpanel also participate in the DNB household survey that is run by CentERdata as well. This survey gathers information on the financial situation and investment choices of the households. The sample is updated semi-annually with new panel members to keep the sample representative of the Dutch population. ‘‘The collection and availability of a wide set of psychological factors makes the DHS data set unique and particularly suited to studying individual preferences and financial choices.’’ (Teppa and Vis 2012). As evidenced by multiple publications in scientific (e.g. Gaudecker (forthcoming), Van Rooij et al. (2011a,b)) and practitioner journals.[49]

The Dutch pension system[50] provides an ideal situation to investigate to what extent values drive financial decision-making. Since the Dutch pension system up to date is a defined benefit (DB) culture the individuals do not have to take riskiness into account, only the final payments matter. We define a setting in which we ask respondents how much they value typically applied socially responsible investment screens. In the next step we ask the respondents if they are willing to give up a small part of their pension for the application of these screens. We state explicitly that it will cost them money, reflecting the direct costs of screening and the sacrificed return from reducing the investable universe (Hong & Kacperczyk 2009). Because we assume pension fund managers maximize returns within the boundaries of their mandates, screening can never lead to an increase in expected financial returns. This line of questioning provides us with a way to test if beneficiaries are willing to give up some of their wealth in order to (partially) align the pension investments with their

[49] www.centerdata.nl/en/publications

[50] For an elaborate description of the Dutch pension system see Van Rooij et al. (2007).

values. Alternatively stated, we test the disutility from investments in undesirable financial products.

We focus on stated preferences towards social and environmental screening. Since we investigate a financial product we avoid the difficulties that a lot of studies into ethical products have regarding longer travel distance and availability of the products (e.g. Laroche et al. 2001, Becchetti and Rosati 2007).

The main theme of the questionnaire is sustainable behavior of households, and was sent to a representative sample of the Dutch society by CentERdata at the fourth of March 2011. The respondents were given until eight March 2011 to answer the questions. While the response rate concerning the entire survey was 63% (1843 out of 2878 members) we focus in this study on those household members of at least 20 years of age[51] (1766 members). The average participant is a little over 55 years of age, the average household has a total household net income that is €2837 per month, and 54% is male. We also obtain information on drinking and smoking behavior. 18% are smokers, and 27% never drink alcoholic beverages.

Furthermore, we asked the respondents to self asses their risk tolerance when it comes to decision making in the pension domain. In addition to self assessed risk tolerance we use another measure introduced by Barsky et al. (1997). This measure uses forced choices between gambles over a lifetime income to measure risk tolerance. The correlation between the two risk tolerance measures is only slightly above 20% therefore we follow Van Rooij et al. (2007) and use both measures as control variables in our analyses since they apparently measure different dimensions of risk tolerance.

[51] We also repeat all our analyses on a sample of individuals who have at least 20 years of age and cannot be older than 65, which is the legal retirement age. As a result, this sample consists out of 926 participants. All reported results are robust to using this sample.

Besides risk tolerance, the household members estimate their own level of financial sophistication. Around 6% of the respondents claim to be in the two highest categories of financial expertise. However, self assessed financial literacy might not be optimal so we match our original data to a questionnaire that covered financial literacy sent out by CentERdata in May 2011. We are able to match roughly 77% of all (1766) observations. The variables we use include the three basic questions on financial literacy originally designed by Lusardi & Mitchell (2008) expanded with two additional questions on mortgage rents and the relation between interest rates and bond prices[52]. Only 15.4% of the respondents were able to answer all financial literacy questions correctly.

4.4. Results

4.4.1. Measuring social responsibility; Exclusionary strategies

We use several approaches to measure the extent to which people value social responsibility in their pension investments. For the groundwork we use typically applied exclusionary strategies from socially responsible investment practices. The companies excluded operate in the so-called "sin" industries, i.e., alcohol, tobacco, and gambling. Companies operating in these industries are often excluded from the investment portfolios of large institutions that are subject to social norms in the U.S. (Hong and Kacperczyk 2009). The other exclusions are companies that (in)-directly violate human rights or operate in the nuclear energy, weapons manufacturing[53], or pornography industries, since these issues have been increasingly receiving attention in the European institutional investment environment in recent years. This

[52] For exact specification of the financial literacy questions see Appendix B.

[53] Not cluster bombs and nuclear weapons since they are already excluded by the Dutch pension funds.

concept of excluding assets from the investable universe is referred to as exclusionary screening. We let the respondents rate the exclusionary screens on a seven point Likert scale from "very unimportant" to "very important".

From panel A in Table 4.1 we find, on average, respondents deem the exclusion of human rights offenders and the weapon industry as more important than other types of exclusionary screens. On average households seem to care the least about screening investments in the alcohol industry. Interestingly, the lowest ranked screens show the highest standard deviation, which means that respondents' values tend to be more consistent for exclusionary screens which were more highly ranked on average. Applying these investment screens is one approach to capture the values of a larger part of society. However, before drawing inferences from these statistics we want to verify if the reported values make sense. We do this by comparing the stated preferences to the behavior of the respondents.

Table 4.1 Importance of exclusionary screens

This table summarizes the self reported importance level of several exclusionary screens often applied in practice by (social) investment funds. The answers range from 1=very unimportant to 7=very important. Panel A contains basic summary statistics. Panel B until D compare the answers over different groups of respondents testing the uncontrolled mean differences based on gender, smoking, and drinking behavior. Panel reports on the importance of the alcohol exclusionary screen and drinking behavior. *, **, *** represent significance at the 10%, 5%, and 1% respectively.

Summary statistics for the Exclusionary screens

Panel A

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Weapons	1766	5.58	1.93	1	7
Alcohol	1766	4.11	1.90	1	7
Tobacco	1766	4.54	1.94	1	7
Gambling	1766	4.95	1.97	1	7
Sexind	1766	5.10	2.02	1	7
Nuclear_en	1766	4.52	2.12	1	7
Human	1766	5.89	1.77	1	7

Table 4.1 Continued

Panel B

	Female (815 obs.)		Male (951 obs.)		Difference		Tests
Variable	Mean	Std. Dev.	Mean	Std. Dev.			t-stat
Weapons	5.88	1.76	5.32	2.03	0.56	***	6.14
Alcohol	4.44	1.86	3.83	1.89	0.62	***	6.88
Tobacco	4.77	1.89	4.35	1.97	0.42	***	4.57
Gambling	5.25	1.88	4.70	2.01	0.56	***	5.98
Sexind	5.53	1.88	4.73	2.07	0.80	***	8.50
Nuclear_en	5.08	1.97	4.04	2.12	1.04	***	10.63
Human	6.08	1.66	5.74	1.85	0.34	***	4.01

Panel C

	Smoker (316 obs.)		Non-smoker (1450 obs.)		Difference		Tests
Variable	Mean	Std. Dev.	Mean	Std. Dev.			t-stat
Weapons	5.47	1.96	5.60	1.92	-0.14		-1.14
Alcohol	3.71	1.82	4.20	1.91	-0.48	***	-4.13
Tobacco	3.82	1.77	4.70	1.95	-0.88	***	-7.44
Gambling	4.75	1.97	5.00	1.97	-0.24	**	-1.98
Sexind	4.97	2.01	5.13	2.03	-0.16		-1.26
Nuclear_en	4.37	2.11	4.55	2.12	-0.18		-1.37
Human	5.82	1.77	5.91	1.77	-0.09		-0.86

Panel D

	Drinker (1285 obs.)		Non-drinker (481 obs.)		Difference		Tests
Variable	Mean	Std. Dev.	Mean	Std. Dev.			t-stat
Weapons	5.62	1.90	5.48	2.01	0.14		1.35
Alcohol	3.97	1.86	4.49	1.95	-0.53	***	-5.21
Tobacco	4.50	1.94	4.65	1.96	-0.15		-1.46
Gambling	4.97	1.95	4.90	2.03	0.07		0.73
Sexind	5.09	1.99	5.12	2.12	-0.04		-0.37
Nuclear_en	4.42	2.11	4.79	2.11	-0.37	***	-3.30
Human	5.93	1.74	5.80	1.85	0.13		1.36

Panel E

Units of alcohol per week	Obs.	Percent	Mean	Std. Dev.	Min	Max
None	481	27%	4.49	1.95	1	7
1 to 5	705	40%	4.12	1.88	1	7
6 to 10	353	20%	3.94	1.86	1	7
11 to 20	189	11%	3.68	1.81	1	7
>20	38	2%	2.92	1.40	1	5

As alluded in our introduction, we expect smokers and drinkers to have different values towards the exclusion of the respective industries from their pension investments. Panel C of Table 4.1 displays the results of differences in mean values with the accompanying test statistics for the tobacco industry screen between smokers and non-smokers. The difference is

negative and highly statistically significant with a t-statistic of -7.44. It is also a natural step to compare how the alcohol exclusionary screen is valued by drinkers (people who consume at least one glass of alcohol a week) and non-drinkers. Again, the difference in importance is statistically significant with a t-statistic of -5.21 (see panel D of Table 4.1). More specifically, for drinkers we observe that the mean values people attribute to the alcohol industry exclusionary screen gradually decrease in the amount of alcoholic beverages consumed per week (see Table 4.1 panel E). We interpret this as evidence that peoples' reported attitudes correspond with their day-to-day behavior. In addition, similar to studies of consumer products (e.g. Auger et al. 2008, De Pelsmacker et al. 2005), we find significant gender differences. On average women account higher values to all exclusion criteria than do men (see panel B of Table 4.1). In sum, Table 4.1 shows that there is significant (between groups) variation in the attitudes towards environmental and social screens in pension investments, which is in line with our first hypothesis.

4.4.2 Are pension investment decisions influenced by social investment preferences?

Our survey results show conclusive evidence of variation in the valuation of several social investment strategies. We now turn to focus on financial decisions people make and specifically if those decisions change when offered a more socially responsible alternative. It is also important to check if the respondents are able to make financial decisions while simultaneously taking their non-financial preferences into account.

4.4.2.1. Positions in stocks vs. bonds

At the start of the survey the respondents fill out what part of their pension allowance they would like to invest in stocks given that currently the average pension fund invests 40% of

the investment in stocks and 60% in bonds. For the sake of simplicity we refrained from considering other types of investments following Van Rooij at al. (2007). To overcome confusion about stocks and bonds we state that bonds are characterized as low risk, low expected return investments and stocks as high risk, high expected return investments[54].

In a next step we introduce a stock portfolio that has exactly the same characteristics as the current portfolio the pension fund holds for the participant. It offers the same payment and is exposed to the same amount of risk, but applies all the described exclusionary screens to the investment portfolio. We ask if they would like to invest more, the same, or less if the basket of stocks they are offered is socially screened.

In panel A of Table 4.2 we show that 17.5% of our respondents answer this question with "less". This result suggests that if all respondents were rational agents, at least this 17.5% derive *negative* utility from excluding the controversial industries because the screens are applied with ceteris paribus conditions.[55] Nevertheless, the sum of stated preferences towards the exclusionary screens for this group is *higher* than for the beneficiaries who want to invest an equal amount given the screened portfolio. Therefore, it is plausible that on average these beneficiaries experience problems in making financial decisions. In support of this suggestion, unreported tests show that this group has significantly lower financial expertise and a lower level of education than other respondents.

On average, people who want to invest more in stocks given the portfolio is screened have significantly stronger stated preferences towards the exclusionary screens (Table 4.2 panel A). Put differently, they assess higher values of importance to the exclusionary screens than all other respondents. In unreported analyses we verify that this group consists out of

[54] Van Rooij et al. (2011b) reports that around 40% of household members do not know the difference between the risk-return characteristics of stocks and bonds.

[55] *At least* 17.5% since risk aversion might withhold beneficiaries from switching their holdings.

individuals who are more risk tolerant therefore they are more eager to shift their portfolio to more stockholdings.

Table 4.2 Financial choices and screening valuation, three candidates.

This table summarizes the sums of self reported importance level of screens (often applied in practice by social investment funds) sorted by three financial choice variables. The answers range from 1=very unimportant to 7=very important per screen. Panel A compares the answers between the groups that want to invest less, the same, or more in stocks if the portfolio applies the exclusionary screens. Panel B does the same for portfolio preferences "screened", "indifferent", and "conventional". Panel C reports the differences of the self reported important levels between the group willing to pay for screening and the group that is not as well as summary statistics on the willingness to pay variable. This is done using two different definitions of the WTP variable. In each panel we test difference between the outer categories and report the two tailed t-statistics. Standard deviations are reported in parentheses. *, **, *** represent significance at the 10%, 5%, and 1% respectively.

Panel A: % of stock invested

	More	The same	Less	More-Less	Tests (t-stat)
Sum excl. screens	38.20	34.13	35.19	3.01 ***	2.62
Std. Dev.	9.26	10.69	12.57		
Observations	153	1304	309		
Percentage	8.66%	73.84%	17.50%		

Panel B: Preferred portfolio

	Screened	Indifferent	Conventional	Screen.-Conv.	Tests (t-stat)
Sum excl. screens	36.91	30.47	33.91	3.00 ***	4.85
Std. Dev.	9.31	12.59	11.35		
Observations	964	451	351		
Percentage	54.59%	25.54%	19.88%		

Panel C: Willingness to pay (WTP)

WTP(a)	Yes = 5-7	Neutral = 4	No = 1-3	Yes-No	Tests (t-stat)
Sum excl. screens	37.57	34.6	30.96	6.61 ***	11.49
Std. Dev.	9.13	9.68	12.60		
Observations	805	336	625		
Percentage	45.58%	19.03%	35.39%		

WTP(b)	Yes = 6-7	Neutral = 3-5	No = 1-2	Yes-No	Tests (t-stat)
Sum excl. screens	38.44	34.59	30.89	7.55 ***	10.38
Std. Dev.	8.94	9.70	13.36		
Observations	494	795	477		
Percentage	27.97%	45.02%	27.01%		

The most important result from panel A in Table 4.2 is that almost all participants answer “the same amount of stocks” meaning that they are indifferent between their current holdings and the holdings of the screened portfolio. Another interpretation of this result is that they have such high levels of risk aversion that the gain in utility is not large enough for them to switch to a slightly riskier portfolio (more stock holdings).

Summarizing, these findings suggest that financial il- literacy is prevalent as a large group of the respondents report questionable financial choices given their attitudes towards social screening. We can see this from the reluctance of respondents to change hypothetical stock holdings and from difference tests in stated screening preferences.

4.4.2.2. Preference between stock baskets

Analyzing percentages invested in hypothetical stock portfolios seems not to be the optimal way to investigate the effects of values on financial decisions. Therefore we also included a more simple measure in which we ask the respondents directly for their preferences between stock baskets. A basket that is equal to the current portfolio of their pension fund and one that applies the exclusionary screens, holding all characteristics like risk and return equal. Almost 55% of the respondents prefer the screened basket of stocks given equal characteristics. In panel B of Table 4.2 we show that these respondents account a higher importance level to the sum of all exclusionary screens than all other participants on average.

Other respondents prefer their conventional portfolio or do not have a preference. The difference in importance levels between those two groups points into the opposite direction from what we would expect (the group that is indifferent reports lower values than the group that prefers the conventional portfolio). In an unreported test we verify that this difference is

significant at the 1% level. Again these results hint towards an explanation in which a large part of the participants are not able to make these simplistic financial decisions. Besides that we take from this question that the majority of the respondents positively value social responsibility in their pension decisions as 55% prefer the screened portfolio.

4.4.2.3. Willingness to pay

We also measure if the values of the beneficiaries towards the screening criteria translate into a willingness to pay (henceforth WTP) for these values. Put differently, is the net gain in utility (if any) enough to lead to the participants accepting a lower pension income? In defined benefit pension plans the pension funds have the duty to invest the allowances for participants. This makes it possible to extract values from the answers the respondents give because they do not have to take riskiness into account. We explicitly explain to the respondents that only their monthly pension entitlements will vary.

The respondents rate their willingness to pay on a seven point Likert scale that ranges from "No, certainly not" to "Yes, certainly". From these seven point scales we create two indicators of WTP, WTP(a) and (b), with the outcomes "Yes", "Neutral", and "No". For WTP(a) we assign "Neutral" if the value is 4, and "Yes" ("No") if the value is higher (lower) than 4. For WTP(b) we are more conservative by assigning "Neutral" if the value is 3, 4, or 5. The results can be seen in panel C of Table 4.2. Beneficiaries who are willing to give up pension income have the highest stated preferences towards the proposed screens compared to the other groups. The differences between the beneficiaries who want to pay and the ones who do not are statistically significant with t-statistics higher than 10 for both the conventional WTP(a) and the more conservative WTP(b).

4.4.3. Results: Financial literacy and translating values into sound financial decisions[56]

From Section 4.4.2 it follows that the three financial choices made do not always match with the behavior of a rational agent. On the one hand this is to be expected since previous research reports low levels of financial literacy among household members (e.g. Van Rooij et al. 2011a,b). On the other hand we posed the questions in such a way that only very limited financial knowledge is needed to answer them. Therefore this section takes a deeper look into the relationship between financial decision making consistent with economically rational behavior and financial literacy.

A first glance at the answers shows that it is difficult to explain the answers to the different questions using economic rationale. For instance, the significant variation in the self reported preference levels of the applied exclusionary screens between groups leads us to expect different in- vestment preferences when these screens are applied. In contrast we see that the vast majority of individuals make the same hypothetical investment decision. Therefore, to relate the answers of the percentage invested in stocks question to the stated preferences towards the social screens, individuals should have an extreme aversion against in- vesting in stock market for example due to risk aversion or distrust (e.g. Guiso et al. 2008).[57] Another more plausible explanation is that beneficiaries are simply not able to translate their preferences into a rational utility maximizing financial choice when at the same time taking risk and return into account. Remember, that the other two financial choice measures (preferred basket and WTP) are not influenced by risk aversion in this setting.

[56] All multivariate results reported in the remainder of this paper are robust to using category dummies instead of ordered variables for education, age, and income.

[57] The disutility from investing in the stock market should be very high since the beneficiaries can change the stockholdings with 1% increments.

To measure if the respondents make rationally consistent financial choices we create four different dummy variables that measure if the financial choices of beneficiaries are consistent with rational behavior.[58] We assume a very general additive utility function where agents get positive utility from expected returns, negative utility from risk, and neutral/positive/negative utility from exposure to the social and environmental attributes of the portfolio of stocks (the screens in our setting). Therefore, the consistent answers can be summarized as follows: the type whose utility is not influenced by social and environmental screens should have no preference between the screened and unscreened portfolio, he should not be willing to pay and the percentage invested in stocks should be the same under the screened and unscreened portfolio situations since the screens were applied with ceteris paribus conditions. We perform this exercise for all three rational agent types in Appendix D. When the stated financial choices are not in line with any of the three types, the choices are inconsistent with rational behavior. We define Inconsistency1 to be one if the answers to the percentage invested in stocks and the WTP(b) question are not consistent with rational behavior (e.g. ‘‘I want to invest less in stocks’’ and ‘‘yes, I am willing to pay for screening’’).[59] Inconsistency2 is one if the answers to the percentage invested in stocks and the preferred basket question are not consistent with rational behavior (e.g. ‘‘I want to invest less in stocks’’ and ‘‘I prefer the socially screened basket’’). Inconsistency3 is one if the answers to the WTP(b) and the preferred basket question are not consistent with rational behavior (e.g. ‘‘I have no preference between baskets and yes, I am willing to pay for screening). Finally, Inconsistency4 is one if at least one inconsistency from rational mean–variance behavior is observed.

[58] See appendix D for the exact description of the how the Inconsistency variables are defined.

[59] All our results are robust to using the less conservative WTP(a) to create the ERROR variables.

Table 4.3 Financial literacy and making consistent financial choices

This table presents marginal effects measured at mean values after a probit estimation on four different Inconsistency dummies that take on a value of one if the respondent makes a choice that is not consistent with the previously made financial choice. Inconsistency1 measures the consistency of answers between the ""% of stock invested" and "WTP(b)", Inconsistency2 between "% of stock invested" and "preferred portfolio", Inconsistency3 between "WTP(b)", and "preferred portfolio", Inconsistency4 measures is only 0 for those who make consistent choices for all three variables. See appendix D for the exact specification of the Inconsistency measures. Panel A and B use the sum of the correct answers to the financial literacy questions as independent variables. Panel C and D include a factoring method explained in detail in Appendix C. In the first step of the two-step regression model the standard errors are clustered by household. We report the R-squared of the first stage. Z-statistics are in parentheses. *, **, *** represent significance at the 10%, 5%, and 1% respectively.

	Inconsistency1 11%	Inconsistency2 15%	Inconsistency3 26%	Inconsistency4 37%
Panel A: Without controls				
Finlitsum	-0.019**	-0.042***	-0.008	-0.048***
	(-2.158)	(-4.110)	(-0.621)	(-3.366)
1st stage pseudo-R2	0.005	0.015	0.000	0.007
Test Fin. Literacy=0	4.654	16.920	0.385	11.330
P-value	0.031	0.000	0.535	0.001
Panel B: With controls				
Finlitsum	-0.012	-0.029***	-0.008	-0.037**
	(-1.377)	(-2.702)	(-0.598)	(-2.389)
l_hhnetincome	0.003	-0.009	-0.016*	-0.024**
	(0.406)	(-1.160)	(-1.704)	(-2.277)
Education	-0.015**	-0.019***	-0.022**	-0.034***
	(-2.548)	(-2.831)	(-2.498)	(-3.577)
Age	0.002***	0.002**	0.002*	0.004***
	(3.153)	(2.351)	(1.878)	(3.382)
Rural	0.005	0.006	0.010	0.008
	(0.772)	(0.742)	(1.033)	(0.751)
Hhsize	-0.014	0.005	0.012	0.009
	(-1.535)	(0.460)	(0.987)	(0.704)
Male	0.008	0.002	0.059**	0.065**
	(0.459)	(0.083)	(2.388)	(2.353)
Dsmoker	-0.011	-0.025	-0.003	-0.014
	(-0.513)	(-1.002)	(-0.086)	(-0.370)
Drinker	-0.002	-0.006	0.013	-0.005
	(-0.218)	(-0.603)	(1.101)	(-0.381)
Finexpert_self	-0.003	-0.011	0.003	-0.005
	(-0.565)	(-1.608)	(0.355)	(-0.544)
Risktol_Barsky	0.001	0.004	0.005	0.006
	(0.160)	(0.573)	(0.530)	(0.558)
Risktol_self	0.007	0.006	0.030***	0.028**
	(0.994)	(0.694)	(2.904)	(2.412)
l_time	0.010	0.011	0.009	0.016
	(0.840)	(0.856)	(0.513)	(0.787)
1st stage pseudo-R2	0.047	0.044	0.026	0.038
Test Fin. Literacy=0	1.903	7.360	0.358	5.703
P-value	0.168	0.007	0.550	0.017

Table 4.3 continued

	Inconsistency1	Inconsistency2	Inconsistency3	Inconsistency4
Panel C: Factor variables without other controls				
Factor1	-0.013*	-0.019**	-0.011	-0.033**
	(-1.705)	(-2.137)	(-0.957)	(-2.454)
Factor2	-0.011	-0.035***	-0.001	-0.032**
	(-1.279)	(-3.501)	(-0.077)	(-2.399)
1st stage pseudo-R2	0.005	0.015	0.001	0.007
Test Fin. Literacy=0	4.727	16.790	0.920	11.750
P-value	0.094	0.000	0.631	0.003
Panel D: Factor variables with other controls				
Factor1	-0.007	-0.009	-0.004	-0.017
	(-0.885)	(-0.996)	(-0.309)	(-1.244)
Factor2	-0.008	-0.028***	-0.006	-0.031**
	(-1.018)	(-2.750)	(-0.496)	(-2.197)
l_hhnetincome	0.003	-0.009	-0.016*	-0.024**
	(0.404)	(-1.134)	(-1.702)	(-2.275)
Education	-0.015**	-0.020***	-0.022**	-0.035***
	(-2.540)	(-2.969)	(-2.499)	(-3.613)
Age	0.002***	0.002***	0.002*	0.004***
	(3.152)	(2.398)	(1.876)	(3.391)
Rural	0.005	0.005	0.010	0.008
	(0.753)	(0.650)	(1.020)	(0.706)
Hhsize	-0.014	0.005	0.012	0.009
	(-1.537)	(0.482)	(0.986)	(0.704)
Male	0.008	0.004	0.060**	0.067**
	(0.458)	(0.192)	(2.385)	(2.390)
Dsmoker	-0.011	-0.025	-0.003	-0.014
	(-0.512)	(-1.011)	(-0.084)	(-0.373)
Drinker	-0.002	-0.006	0.013	-0.005
	(-0.219)	(-0.631)	(1.099)	(-0.382)
Finexpert_self	-0.004	-0.012*	0.003	-0.006
	(-0.596)	(-1.654)	(0.342)	(-0.587)
Risktol_Barsky	0.001	0.005	0.005	0.006
	(0.167)	(0.669)	(0.535)	(0.592)
Risktol_self	0.007	0.006	0.030***	0.028**
	(0.991)	(0.698)	(2.908)	(2.414)
l_time	0.010	0.010	0.009	0.016
	(0.840)	(0.773)	(0.504)	(0.758)
1st stage pseudo-R2	0.046	0.045	0.026	0.038
Test Fin. Literacy=0	1.781	8.281	0.330	6.097
P-value	0.410	0.016	0.848	0.047
Observations	1368 (1363 with controls)			

This results in four variables that measure the consistency of financial decisions in different ways. The correlation is highest among the first and the second in- consistency (0.58), and lowest between the second and the third inconsistency (0.01). This implies that we

are not measuring the same thing in three different ways. 11% of the respondents make the first inconsistency, compared to 15%, 26% and 37% reporting the second, third and the fourth inconsistency respectively. These results show that a considerable part of our sample is not able to consistently incorporate their preferences into their financial decisions in different situations, confirming our third hypothesis.

Furthermore, to measure if the financial choices are related to financial literacy we use the expanded literacy questions from Lusardi and Mitchell (2008). Using the answers to five questions we create two different literacy indexes. The first one is simply the sum of all correct answers to the five literacy questions (Finlitsum). The second index is created using a principal component factoring analysis retaining only the components with an eigenvalue greater than one (Factor1 and Factor2).[60]

To test our hypothesis that less financially sophisticated individuals are less capable of making financial decisions while taking social and environmental preferences into account we estimate several probit models with the In- consistency measures explained by the financial literacy variables defined as Finlitsum, Factor1 and Factor2. In Table 4.3 we report marginal effects estimated at mean values.

For all specifications reported in Table 4.3 we find negative coefficients on the financial literacy variables, which are significant at the 5% level in three out of four specifications. [61] These results strongly support our fourth hypothesis. The model specifications that include control variables have lower loadings on the literacy variables. This is not surprising as education is among the most important drivers of financial literacy (Van Rooij et al. 2011a,b).

[60] See appendix C for more information on the creation of the financial literacy variables.

[61] Financial literacy is not significant in explaining an inconsistent choice between the willingness to pay and the preferred basket questions. It seems that financial illiteracy does a better job at explaining the inconsistencies resulting from a combination with the more ‘‘complicated’’ question. Whereas education is significant in explaining all inconsistency variables.

As expected, the coefficients on education load significantly negative indicating that on average people who en- joyed a higher level of education are less likely to make financial choices that contradict with economically rational behavior. Income has a negative effect while age has a strong positive effect on all four inconsistency dummies. In addition we find that being male and having a higher self assessed risk tolerance increases the likelihood of making the third and fourth inconsistency.

Summarizing, we have shown that over one third of the beneficiaries do not incorporate their preferences into their financial decisions consistent with rational behavior and that a lack of financial sophistication can partially ex- plain this. These findings have important consequences, if the fiduciaries want to take the preferences of beneficiaries into account they need to be very careful in designing a solution to both the problems of heterogeneous preferences and the low financial literacy of the households. In addition, governments who want to give their population more financial responsibility should also take great care to make sure that individuals are well prepared with good skills and the tools to make financial decisions.

4.5. Additional analyses

In this section we provide further evidence that Dutch pension beneficiaries have positive preferences to several socially responsible investment screens.

4.5.1. Measuring social responsibility; Best practices

Taking the same approach as for the exclusionary screens we ask the respondents to value certain best practices as selection criteria for stocks. Similar to most socially responsible mutual funds we do not explain the details of the criteria as these are rather subjective, and

difficult for respondents to interpret. We only provide the basic idea of the screening process explaining that the companies that are selected perform above average on that practice. This gives us the opportunity to refrain from the numerous possibilities to apply best practices screening. We argue that this is an issue on the implementation of the screens which suits investors who are willing to use these types of investment screens to maximize returns while we investigate if individuals' decisions are driven by their non-financial preferences.

Table 4 panel A shows that from the possible best practices criteria, participants rate employee relations (e.g. pension, health and safety, schooling, anti-discrimination, and work atmosphere) as most important and with the lowest deviation in their responses. While selecting firms based on their charity policy is rated as least important. This is striking since the Dutch gave around €4.3 billion to charities in 2007 which is 0.8% of the GDP (Schuyt et al. 2009). This might have several reasons. Some charities have suffered from bad press due to high salaries paid to their directors. Another reason can be that people prefer to choose themselves which charities to support.

For completeness we compare the average stated preferences towards these screens over different groups, as we did in Table 1 (see panels B to D of Table 4.4). Most striking is that (again) women value all screening criteria significantly higher than men, except for the profit screen. We can conclude that female beneficiaries have a fixed positive effect towards social and environmental screening.

For completeness we compare the average stated preferences towards these screens over different groups as we did in Table 4.1 (see panels B to D of Table 4.4). Most striking is that (again) women value all screening criteria significantly higher than do men, except for the profit screen. Therefore, female beneficiaries have a fixed positive effect towards *social and environmental* screening.

Table 4.4 Importance of best practices screens

This table summarizes the self reported importance level of several best practices screens often applied in practice by (social) investment funds. The answers range from 1=very unimportant to 7=very important. Panel A contains basic summary statistics. Panel B until D compare the answers over different groups of respondents testing the uncontrolled mean differences based on gender, smoking, and drinking behavior. *, **, *** represent significance at the 10%, 5%, and 1% respectively.

Summary statistics for the Positive screens

Panel A

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Recycling	1766	5.53	1.43	1	7
CO2	1766	5.44	1.49	1	7
Employees	1766	5.76	1.37	1	7
Community	1766	5.33	1.46	1	7
Charity	1766	4.82	1.62	1	7
Profit	1766	5.15	1.44	1	7

Panel B

	Female (815 obs.)		Male (951 obs.)		Difference		Tests
Variable	Mean	Std. Dev.	Mean	Std. Dev.			t-stat
Recycling	5.66	1.37	5.43	1.48	0.22	***	3.29
CO2	5.63	1.41	5.28	1.54	0.35	***	4.99
Employees	5.94	1.28	5.61	1.42	0.33	***	5.11
Community	5.47	1.41	5.20	1.49	0.27	***	3.93
Charity	4.99	1.56	4.67	1.67	0.32	***	4.14
Profit	5.21	1.42	5.11	1.45	0.10		1.44

Panel C

	Smoker (316 obs.)		Non-smoker (1450 obs.)		Difference		Tests
Variable	Mean	Std. Dev.	Mean	Std. Dev.			t-stat
Recycling	5.30	1.48	5.59	1.42	-0.29	***	-3.25
CO2	5.26	1.51	5.48	1.48	-0.22	**	-2.38
Employees	5.62	1.43	5.79	1.35	-0.17	**	-2.02
Community	5.15	1.56	5.37	1.43	-0.21	**	-2.35
Charity	4.65	1.72	4.85	1.60	-0.20	**	-2.02
Profit	5.12	1.49	5.16	1.43	-0.05		-0.52

Panel D

	Drinker (1285 obs.)		Non-drinker (481 obs.)		Difference		Tests
Variable	Mean	Std. Dev.	Mean	Std. Dev.			t-stat
Recycling	5.58	1.40	5.41	1.53	0.16	**	2.15
CO2	5.45	1.49	5.41	1.49	0.04		0.50
Employees	5.75	1.36	5.79	1.39	-0.04		-0.54
Community	5.30	1.45	5.39	1.47	-0.09		-1.14
Charity	4.79	1.62	4.89	1.65	-0.10		-1.19
Profit	5.18	1.43	5.10	1.48	0.08		0.98

4.5.2. Are beneficiaries willing to pay?

In total we now have two sets of social investment screens and the stated preferences towards these screens. We continue by investigating if beneficiaries actually get positive utility from investing their pension endowments more responsibly. We do this by asking the respondents if they are willing to sacrifice pension income for the application of the screens that they valued higher than 4 out of 7. When no screen was ranked higher than 4 we select all criteria (this happens in 38% of all cases). We do this for both the exclusionary screens presented in Table 4.1 and the best practices screens presented in Table 4.4.

In two steps the respondents had to fill out if they would accept a monthly pension entitlement that is lower than their expected entitlement they receive under the current investment policy. Panel A of Table 5 shows that 25.72% (28.71%) do not want to give up anything or only a negligible part of their pension for the exclusionary (best practices) screens that they personally scored earlier in the survey. Roughly 40% of the respondents are willing to give up 5% of their pension income after retirement. Since it is well known that framing can influence the outcome of the results with these types of questions we refrain from interpreting these percentages too closely.[62] These numbers do inform us that we can confirm our second hypothesis that individuals derive positive utility from the social responsibility attribute since over 70% of our respondents claims to be willing to give up a considerable part of their pension income for such an alignment.

[62] The percentages are chosen in consensus with the CentERdata staff. More realistic (lower) percentages would cause almost all participants to declare themselves willing to pay since they perceive it as a negligible amount.

4.5.3. Determinants of willingness to pay

To analyze willingness to pay we generate a dummy variable that takes on the value of 1 if the respondent is willing to give up pension income for (a partial) alignment of pension investments with his social and environmental preferences. First we verify that the beneficiaries who rate the screens as more important are also more likely to be willing to pay for the implication of these screens (Table 4.5 panel B). The sum of attitude scores for all exclusionary screens is significantly higher (t-statistic of 10.49) among the respondents who want to give up part of their pension income. Also the number of screens rated higher than 4 is significantly higher. In panel C of Table 4.5 we verify these results for the best practices screens. Moreover, the correlation between the dummy on willingness to pay for exclusionary screens and the dummy for best practices criteria is 85% implying that the respondents who get positive utility from screening obtain this for different types of social screening.

In Table 4.6 we make explicit for which type of person the surplus utility of screening is high enough to overcome any pecuniary motives. Therefore we use the WTP(p) dummy in a probit regression framework with the log of net monthly household income, gender, and some other controls often used in household finance studies (e.g. Van Rooij et al. 2007, Van Rooij et al. 2011a,b, Renneboog and Spaenjers 2011). In addition we add a dummy that takes on the value of one if the respondent scores at least one screen higher than four out of seven (Dimportant). Finally, we add risk tolerance and financial expertise variables. The first 4 columns of Table 4.6 show the marginal effects predicted at the mean value. Individuals who enjoyed a higher level of education and individuals from higher income categories are significantly more likely to be willing to pay for personalized social screening. The most significant result is that individuals who have a positive stated preference towards at least one of the social screens are 40% more likely to sacrifice pension income. This finding confirms

results from studies on consumer products that find a higher likelihood of willingness to pay for non-product attributes when a positive attitude to that attribute is displayed (Auger et al. 2003, 2008). Our results are robust to excluding respondents who make inconsistent financial choices and in unreported analyses we show that excluding measures of financial expertise and risk tolerance or using different variables to measure preferences does not materially alter the results.

Table 4.5 Willingness to pay for personalized social pension investments

This table reports the willingness to pay for personalized social pension investment screening. We personalize the investments by selecting only those screens valued higher than four on a seven point Likert scale. If none of the screens is valued higher than four we select all screens. Panel A reports the fractions of the sample and the accompanying answers. Panel B and C report the average sum of values sorted by willingness to pay and the number of screens valued higher than 4. The differences and two tailed t-statistics are reported in the most right columns. *, **, *** represent significance at the 10%, 5%, and 1% respectively.

Panel A: WTP(p) variable

	Exclusionary screens		Best practices screens		
WTP	Obs.	Percent	Obs.	Percent	WTP(p)
<1%	456	25.72%	509	28.71%	No
1%	84	4.74%	113	6.37%	Yes
2%	446	25.16%	460	25.94%	Yes
≥ 5%	787	44.39%	691	38.97%	Yes

Panel B: Assessed values and WTP(p)for Exclusionary screens

WTP(p)	Yes (1317 obs.)		No (456 obs.)		Difference	Tests (t-stat)
	Mean	Std. Dev.	Mean	Std. Dev.		
Sum excl. screens	36.26	(9.60)	30.18	(13.25)	6.08 ***	10.49
#screens >4	4.69	(2.18)	3.11	(2.80)	1.58 ***	10.80

Panel C: Assessed values and WTP(p)for Best practices screens

WTP(p)	Yes (1264 obs.)		No (509 obs.)		Difference	Tests (t-stat)
Sum B.P. screens	33.08	(6.14)	29.42	(8.63)	3.66 ***	10.04
#screens >4	4.84	(1.54)	3.70	(2.36)	1.14 ***	10.45

The effects are almost identical for the best practices screens (see the final columns in Table 4.6). Interestingly, drinker and male are both significant, which decreases the probability of WTP by 6% for men and for drinkers by 2%–3% per increment. The gender effect is in line with Andreoni and Vesterlund (2001) who find that females are more willing to engage in altruism than do men when altruism is expensive.[63] Also with Laroche et al. (2001) who find that especially women have a higher willingness to pay for environmentally friendly products. Still, by far the most economically and statistically significant determinant is whether or not the beneficiary has a positive stated preference towards social screening. Therefore we test determinants of having this preference and report the results in Table 4.7.[64]

The results from all columns in panel A of Table 4.7 show that female beneficiaries have stronger preferences towards social screening, they are around 4% more likely to score at least one of the screens higher than 4 out of 7. Another interesting result is that individuals with a higher level of education are more likely to have a positive preference- and to appoint the maximum score of 7 out of 7 towards at least one exclusionary screen while at the same time education has no effect on the sum of all screen ratings. All else equal older people rate all screens higher on average. Furthermore, individuals who smoke and are more tolerant to risk are less likely to prefer social screening.

[63] We argue that the amounts of willingness to pay we demand from the individuals are definitely expensive.

[64] We take the approach of two separate probit regressions because the residuals of the two models are not significantly correlated (this correlation coefficient is usually referred to as rho). We used a seemingly unrelated regressions approach adjusted for probit models with standard errors clustered at the household level and adjusted for heteroskedasticity as well as a model with unadjusted standard errors. In the former we use a Wald test (p-value = 0.80) and for the latter a likelihood ratio test (p-value = 0.72). Both tests cannot reject the null of uncorrelated residuals, which supports our methodology.

Looking at panel B of Table 4.7 we see that the results are very similar for the best practices screens. The difference is that education has lower predictive power while self declared financial literacy is now significant. Individuals with higher self reported financial literacy report a lower sum of preferences towards all social screens and have around 3% lower probability per increment to appoint a maximum score to a best practices screen. Across the board the pseudo-R2s and the marginal effects are low meaning that if pension funds want to apply socially responsible investment practices they should reveal the preferences of beneficiaries by direct engagement.

4.6. Discussion

A natural problem in the design of our survey lies in the fact that we do not observe the original pension investments of the pension funds and even if we would be able to get this information we are almost certain that the average individual does not possess this information. Therefore our results can also be interpreted as a lack of transparency of the pension funds' investments. Since beneficiaries who are willing to accept a lower pension for the application of these screens *consider* their pension fund holdings to be less socially responsible than what they would ideally like them to be. This does not necessarily mean that the holdings do in fact not meet the demands of the pension participants to a large degree.

We also asked the respondents if they missed certain screens in the survey. The screens most often mentioned were companies with an excessive bonus culture, companies that engage in animal testing or other business practices that violate animal wellbeing, and using the environmental dimension as an exclusionary screen (worst polluting) rather than a selection screening method. Therefore the percentage of people willing to pay for somewhat customized responsible pension investments might be even larger than reported here.

Table 4.6 Determinants of WTP for personalized pension investments

This table reports the marginal effects estimated at mean values after estimating probit models. The dependent variable is a dummy that is 1 for individuals who are willing to pay for socially screened pension investment portfolios. The individuals are offered a customized choice that selects only those screens rated >4 out of 7. Panel A presents the results for exclusionary screens on the full sample (columns 1 and 2) and a subset which excludes all respondents who make an inconsistent choice (columns 3 and 4). In panel B we do the same for the best practices screens. In addition columns 5 and 6 exclude respondents who only value the "profit" screen higher than 4 out of 7. In the first step of the two-step regression model the standard errors are clustered by household. We report the R-squared of the first stage. Z-statistics are in parentheses. *, **, *** represent significance at the 10%, 5%, and 1% respectively.

	Exclusionary screens				*Best practices screens*					
	Full sample		Conditional Inconsistency4=0		Full sample		Conditional Inconsistency4=0		Excl. only profit>4	
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
Dimportant	0.401***	0.395***	0.455***	0.454***	0.437***	0.422***	0.496***	0.459***	0.444***	0.443***
	(11.249)	(9.402)	(9.382)	(7.924)	(11.422)	(9.279)	(10.138)	(7.677)	(10.806)	(9.075)
l_hhnetincome	0.022*	0.017	0.039**	0.039**	0.044***	0.035**	0.040**	0.042**	0.042***	0.033**
	(1.657)	(1.091)	(2.453)	(2.242)	(3.131)	(2.195)	(2.461)	(2.277)	(2.971)	(2.031)
Education	0.028***	0.021**	0.035***	0.026**	0.033***	0.030***	0.045***	0.043***	0.032***	0.029***
	(3.528)	(2.381)	(3.597)	(2.360)	(4.037)	(3.216)	(4.355)	(3.762)	(3.968)	(3.113)
Age	0.001	0.001	0.001	-0.000	0.001	0.001	0.000	-0.001	0.001	0.001
	(1.465)	(0.855)	(1.121)	(-0.235)	(0.773)	(0.507)	(0.224)	(-0.449)	(0.716)	(0.530)
Rural	0.000	0.003	-0.001	-0.002	0.003	0.004	-0.006	-0.007	0.005	0.006
	(0.025)	(0.264)	(-0.126)	(-0.197)	(0.357)	(0.372)	(-0.525)	(-0.605)	(0.580)	(0.613)
Hhsize	0.006	-0.001	0.007	0.002	-0.000	-0.006	0.009	0.002	-0.001	-0.005
	(0.567)	(-0.094)	(0.505)	(0.155)	(-0.039)	(-0.466)	(0.604)	(0.119)	(-0.045)	(-0.406)
Male	-0.032	-0.050**	-0.036	-0.048	-0.062***	-0.077***	-0.060**	-0.085***	-0.061***	-0.078***
	(-1.493)	(-2.037)	(-1.336)	(-1.562)	(-2.684)	(-3.016)	(-2.080)	(-2.625)	(-2.628)	(-3.045)
Dsmoker	-0.065**	-0.056	-0.022	-0.029	-0.050	-0.055	-0.021	-0.043	-0.050	-0.054
	(-2.214)	(-1.628)	(-0.606)	(-0.703)	(-1.636)	(-1.533)	(-0.559)	(-0.986)	(-1.640)	(-1.495)
Drinker	0.019	0.019	0.026*	0.030*	0.023*	0.019	0.037***	0.037**	0.023*	0.020
	(1.590)	(1.436)	(1.847)	(1.955)	(1.881)	(1.419)	(2.582)	(2.329)	(1.950)	(1.526)
Finexpert_self	0.001	-0.008	0.003	-0.015	0.005	-0.005	0.006	-0.008	0.006	-0.004
	(0.101)	(-0.869)	(0.292)	(-1.369)	(0.638)	(-0.485)	(0.561)	(-0.699)	(0.671)	(-0.446)
Risktol_Barsky	0.024***	0.025***	0.025**	0.033***	0.024***	0.020**	0.018	0.018	0.024***	0.020**
	(2.887)	(2.618)	(2.343)	(2.606)	(2.659)	(2.060)	(1.547)	(1.392)	(2.721)	(2.071)
Risktol_self	0.004	0.002	-0.001	-0.003	0.008	0.009	0.007	0.009	0.007	0.008
	(0.439)	(0.184)	(-0.106)	(-0.194)	(0.738)	(0.756)	(0.532)	(0.606)	(0.705)	(0.709)
Finlitsum		0.065***		0.084***		0.059***		0.065***		0.059***
		(4.494)		(4.490)		(3.749)		(3.373)		(3.734)
Observations	1,764	1,368	1,098	858	1,764	1,368	1,098	858	1,736	1,347
1st stage pseudo-R2	0.107	0.124	0.135	0.175	0.105	0.109	0.137	0.145	0.101	0.109

Table 4.7 Determinants of preference variables for environmental and social screens

This table reports the marginal effects after estimating probit models for columns (a)-(b) and first stage regression coefficients from ordered probit models are shown in columns (e)-(j). The dependent variables used are a dummy for individuals who are willing to pay for socially screened pension investment portfolios (a) and (b), a dummy for individuals that scored at least one screen 7 out of 7 (c) and (d), the sum of the scores for the social screens (e) and (f), quintiles of Total score (g) and (h), and the sum of all exclusionary screens valued higher than 4 out of 7. In panel A we report on the analyses using exclusionary screens, in panel B we report on all the best practices screens except the profit screen. In the first step of the two-step regression model the standard errors are clustered by household. We report the pseudo R-squared of the first stage. Z-statistics are in parentheses. *, **, *** represent significance at the 10%, 5%, and 1% respectively.

Panel A: Exclusionary screens

	Dimportant		1 screen max. score		Total score		Quintiles of total score		Sum of preferred screens	
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
l_hhnetincome	0.021**	0.013	0.022	0.018	0.032	0.020	0.014	0.003	0.035	0.017
	(2.163)	(1.200)	(1.529)	(1.168)	(0.964)	(0.552)	(0.432)	(0.086)	(1.046)	(0.452)
Education	0.017***	0.019***	0.031***	0.033***	0.010	0.031	0.011	0.033	0.031*	0.050**
	(3.097)	(2.942)	(3.753)	(3.540)	(0.530)	(1.511)	(0.584)	(1.555)	(1.689)	(2.443)
Age	-0.000	-0.000	0.000	-0.000	0.008***	0.008***	0.010***	0.010***	0.008***	0.007***
	(-0.072)	(-0.489)	(0.251)	(-0.206)	(3.579)	(3.145)	(4.368)	(3.753)	(3.515)	(2.732)
Rural	-0.007	-0.001	-0.006	0.010	0.002	0.013	0.003	0.013	-0.001	0.009
	(-1.036)	(-0.131)	(-0.613)	(0.921)	(0.101)	(0.601)	(0.150)	(0.556)	(-0.063)	(0.369)
Hhsize	-0.009	-0.010	-0.012	-0.015	0.020	0.015	0.026	0.018	0.008	-0.000
	(-1.105)	(-1.070)	(-1.032)	(-1.173)	(0.815)	(0.545)	(0.991)	(0.604)	(0.294)	(-0.001)
Male	-0.047***	-0.038**	-0.090***	-0.062**	-0.417***	-0.355***	-0.446***	-0.366***	-0.400***	-0.351***
	(-2.812)	(-2.101)	(-3.914)	(-2.406)	(-7.732)	(-6.041)	(-8.097)	(-6.014)	(-7.230)	(-5.742)
Dsmoker	-0.035	-0.060**	-0.027	-0.062*	-0.196***	-0.277***	-0.219***	-0.302***	-0.241***	-0.342***
	(-1.495)	(-2.146)	(-0.882)	(-1.750)	(-3.225)	(-4.074)	(-3.352)	(-4.116)	(-3.578)	(-4.460)
Drinker	0.023***	0.023**	0.020*	0.021	-0.030	-0.019	-0.039	-0.030	-0.014	-0.017
	(2.797)	(2.477)	(1.698)	(1.559)	(-1.176)	(-0.707)	(-1.457)	(-1.031)	(-0.563)	(-0.610)
Finexpert_self	0.002	-0.005	-0.013	-0.025***	-0.039**	-0.047**	-0.040**	-0.052**	-0.018	-0.029
	(0.382)	(-0.779)	(-1.587)	(-2.809)	(-2.193)	(-2.398)	(-2.212)	(-2.537)	(-0.987)	(-1.430)
Risktol_Barsky	0.001	-0.006	0.004	0.002	-0.017	-0.029	-0.019	-0.028	-0.012	-0.032
	(0.103)	(-0.870)	(0.506)	(0.158)	(-0.880)	(-1.320)	(-0.956)	(-1.234)	(-0.561)	(-1.343)
Risktol_self	-0.008	-0.007	-0.040***	-0.035***	-0.046**	-0.042*	-0.057**	-0.047*	-0.049**	-0.043*
	(-1.182)	(-0.865)	(-4.099)	(-3.205)	(-2.100)	(-1.685)	(-2.547)	(-1.808)	(-2.228)	(-1.723)

Table 4.7 Panel A continued

Finlitsum		0.014		0.011		-0.082**		-0.095***		-0.032
		(1.311)		(0.726)		(-2.293)		(-2.712)		(-0.841)
Observations	1,764	1,368	1,764	1,368	1,764	1,368	1,764	1,368	1,764	1,368
1st stage pseudo-R2	0.034	0.038	0.033	0.034	0.011	0.012	0.028	0.029	0.016	0.018

Panel B: Best practices screens

	Dimportant		1 screen max. score		Total score		Quintiles of total score		Sum of preferred screens	
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
l_hhnetincome	0.008	0.007	0.010	-0.003	0.011	-0.004	0.012	-0.008	0.012	-0.000
	(0.859)	(0.684)	(0.649)	(-0.170)	(0.337)	(-0.118)	(0.371)	(-0.202)	(0.324)	(-0.008)
Education	0.010*	0.007	0.002	0.004	-0.010	-0.013	-0.020	-0.027	-0.010	-0.017
	(1.836)	(1.285)	(0.201)	(0.370)	(-0.564)	(-0.653)	(-1.089)	(-1.279)	(-0.501)	(-0.744)
Age	0.002***	0.001	0.007***	0.007***	0.020***	0.018***	0.021***	0.019***	0.017***	0.016***
	(2.578)	(1.371)	(6.804)	(5.136)	(8.511)	(6.663)	(8.563)	(6.738)	(6.749)	(5.279)
Rural	-0.012**	-0.007	-0.020*	-0.018	-0.031	-0.031	-0.027	-0.029	-0.027	-0.025
	(-1.978)	(-1.043)	(-1.885)	(-1.543)	(-1.489)	(-1.376)	(-1.235)	(-1.172)	(-1.145)	(-0.970)
Hhsize	-0.007	-0.013*	-0.012	-0.019	0.006	-0.013	-0.004	-0.020	-0.027	-0.053
	(-1.009)	(-1.778)	(-0.960)	(-1.356)	(0.227)	(-0.424)	(-0.149)	(-0.610)	(-0.971)	(-1.596)
Male	-0.041***	-0.036**	-0.080***	-0.068**	-0.266***	-0.207***	-0.263***	-0.193***	-0.259***	-0.208***
	(-2.777)	(-2.260)	(-3.121)	(-2.310)	(-5.101)	(-3.526)	(-4.877)	(-3.162)	(-4.520)	(-3.198)
Dsmoker	-0.049**	-0.044*	-0.016	-0.040	-0.131**	-0.151*	-0.140**	-0.166**	-0.141*	-0.144*
	(-2.205)	(-1.805)	(-0.499)	(-1.057)	(-2.007)	(-1.930)	(-2.035)	(-2.016)	(-1.890)	(-1.647)
Drinker	0.009	0.007	-0.023*	-0.028*	0.000	-0.006	-0.002	-0.002	0.020	0.026
	(1.076)	(0.853)	(-1.785)	(-1.916)	(0.002)	(-0.183)	(-0.064)	(-0.072)	(0.682)	(0.778)
Finexpert_self	-0.010*	-0.017***	-0.024***	-0.034***	-0.055***	-0.066***	-0.053***	-0.065***	-0.036*	-0.050**
	(-1.918)	(-3.095)	(-2.673)	(-3.404)	(-2.974)	(-3.184)	(-2.773)	(-3.005)	(-1.789)	(-2.229)
Risktol_Barsky	0.002	-0.003	0.021**	0.020**	0.011	-0.003	0.008	-0.003	0.003	-0.014
	(0.378)	(-0.425)	(2.290)	(1.962)	(0.544)	(-0.119)	(0.414)	(-0.139)	(0.142)	(-0.579)
Risktol_self	-0.008	-0.006	-0.051***	-0.048***	-0.075***	-0.071***	-0.074***	-0.072***	-0.046*	-0.039
	(-1.261)	(-0.921)	(-4.549)	(-3.771)	(-3.421)	(-2.845)	(-3.196)	(-2.721)	(-1.921)	(-1.390)
Finlitsum		0.015		-0.008		-0.046		-0.042		-0.010
		(1.612)		(-0.501)		(-1.388)		(-1.212)		(-0.250)
Observations	1,764	1,368	1,764	1,368	1,764	1,368	1,764	1,368	1,764	1,368
1st stage pseudo-R2	0.043	0.052	0.065	0.066	0.021	0.020	0.038	0.038	0.027	0.028

Less of a concern is the fact that the questions reflect simplified versions of reality which is done to partially overcome the low financial literacy of the average house- hold member. A socially responsible investment strategy does not necessarily exclude all assets that do not pass a certain form of screening. Pension funds often interfere with managerial decisions by corporate engagement. In addition, we screened equities whereas debt can also be screened using similar corporate screens or social and political screens on a geographical level. One need only con- sider government bonds of countries where human rights are severely violated. These other types of SRI are ignored in this paper.

Nevertheless, within the current setting of defined benefit pension plans we are able to assess the preferences for socially responsible investing. Our results show a clear opportunity for investment funds to take the heterogeneous nature of individuals into account when designing investment plans. In practice this is reflected by the growth in funds devoted to responsible investment; and is directly measured in our survey data. Pension funds could consider offering a menu of investment choices, which not only include variety in their financial attributes but also in their social responsibility attributes.

4.7. Conclusion

This paper reports on frictions between pension fund participants' social values and the allocation of their pension assets. To our knowledge it is the first empirical study into the effects of social values on pension investment decisions. We provide consistent evidence that beneficiaries do value social responsibility in their pension investments. Companies that operate in the weapons industry and companies that violate human rights amendments are deemed most important to be excluded from the investment portfolios of Dutch beneficiaries in 2011 relative to our proposed screens. While excluding the "sin" industries (alcohol,

tobacco, and gambling) is of a much lower importance to our respondents. It seems to be important for the respondents that companies treat other humans well as the highest rated exclusionary as well as best practices screens are related to human wellbeing.

We further investigated if respondents are able to make financial decisions while simultaneously taking their non- financial attitudes into account and found that over one third reported at least one financial choice inconsistent with rational behavior. In particular, respondents who do not possess the required level of financial literacy to make investment choices and have a lower level of education make choices less in line with rational economic behavior. This is an important finding since the majority of beneficiaries report positive attitudes towards social and environmental screening in their pension investments. Which leads us to consider incorporating non-financial aspects into pension investments.

Summarizing we find that a significant tranche of pension participants do derive positive utility from social screens in their pension investments. If pension funds already try to take into account the preferences of their clientele, they either fail to communicate their social responsibility practices effectively or their social responsibility practices do not meet the preferences of the participants. Because we observe variation in preferences it might not be optimal to provide the pension fund participants with just one pension investment scheme. Our findings have important implications on how socially responsible investment filters into the public domain through a market based system. It is important to reflect upon the influence of financial responsibility in society, whilst reconciling this with the observed level of limited financial literacy, which may stand in the way of providing tailor-made pension investment schemes for all. We leave it to future research how to best tackle these issues.

4.A. Appendix: The questionnaire

The questions on the importance levels of different exclusionary screens

-For each of the exclusionary screens we ask:
What level of importance do you attribute to the following exclusions of your pension investments?

1 = very unimportant
7 = very important

The questions on the importance levels of different best practices screens

-For each of the best practices screens we ask:

The previous was about **not** investing in certain companies, you also have the opportunity to choose in what companies/industries you certainly **do** want to invest in. Think about companies that have a good environmental program, companies that help people in third world countries and companies that take good care of their employees.

What level of importance do you attribute to your pensionfund investing in the following companies?

a. The company recycles a lot
…
f. The company makes sure the profit margins are high (for continuity)

1 = very unimportant
7 = very important

The questions % of stocks invested

Step 1.
What if your employer transfers your pension payments to an individual account after which this money is being invested in stocks and bonds by your pension fund. You have to decide yourself what amount is invested in stocks and what amount in bonds. Stocks have a higher expected return and a higher risk. Bonds have a lower expected return and a very low risk. All Dutch pension funds together held around 40% in stocks and 60% in bonds on average at the end of 2010.
What percentage would you have invested in stocks?
0..100
Step 2.
Intro: Almost all Dutch pension funds invest a part of the/your money in the weapons-*, alcohol-, tobacco-, gambling-,pornography industry, nuclear energy, and companies that (in)direcly violate human rights (like child labor).
*most big pension funds already exclude direct production of nuclear weapons and clusterbombs, however, this industry entails a lot more.

Question: What if your current pension fund decides no longer to invest in previously mentioned industries, keeping all other characteristics like risk and expected return (of the pension funds' investments) equal to the current situation.
Would you like to invest more or less in stocks in this situation?
1 less
2 the same
3 more

The question on the preferred basket of stocks

What if your pension fund offers you the choice to invest your pension in different stock-portfolios:
Package 1 = your current pension fund and stock portfolio
Package 2 = does not invest in previously mentioned industries* but everything else is exactly the same.
*These industries are the weapons-, alcohol-, tobacco-, gambling-,pornography industry, nuclear energy, and companies that (in)direcly violate human rights (like child labor).
Which package do you prefer?
1 Package 1
2 Package 2
3 No preference

The questions on willingness to pay

Imagine that the return on the current pension investments is higher than in the case of the screened investments while the risk is the same. Are you prepared to give up a (very) small part of your pension income for such socially responsible pension composition?
1 = No, certainly not
7 = Yes, certainly

The questions on willingness to pay for a personalized investment

Intro: Imagine that the government decides to switch to a system in which you can choose yourself how much of your pension investments you want to invest in stocks and how much in bonds. Now consider this fund which does not invest in {all the screens the respondent rated higher than 4, if no screen was rated higher than 4 than all screens are applied} has a lower expected payoff but the same risk as your current holdings. We name this portfolio the responsible choice.
Question: If you get the opportunity to invest in the responsible choice. Do you accept a monthly expected pension entitlement that is 2% lower than your current expected pension entitlement? This means that when your current expected pension (including state pension) would be 1100 euros a month that you will agree on a pension that is 1078 euros.

1. *Yes*
2. *No*

If the answer is *yes* we replace the 2% by 5% and ask the question again. The numbers in the example are changed accordingly.
If the answer is *no* we replace the 2% by 1% and ask the question again. The numbers in the example are changed accordingly.
-We repeat this methodology for the best practices screens as well.

4.B. Appendix: Financial literacy questions

1. Do you think that the following statement is true or false? A 15-year mortgage requires higher monthly payments on average than a 30-year mortgage (with equal amount borrowed), but the total rents payed over the whole period are lower for the 15-year mortgage.
2. Imagine that the overall interest rate goes up tomorrow, what will happen to the value of outstanding bonds? (1) They will increase in value (2) They will decrease in value (3) They will have the same value (4) There is no relationship between interest rates and the values of bonds.

4.C. Appendix: Generate financial literacy factors

Table 4.C.1 reports the results of a principal component factor analysis on five variables that measure different aspects of financial literacy. The data is suited for this method since the Bartlett sphericity test returned a p-value of 0.00, and the Kaiser-Meyer-Olkin measure is 0.611. In panel A we report the eigenvalues and the (cumulative) proportions of the variance explained when using 1-5 factors.

Table 4.C.1

Panel A

Component	Eigenvalue	Proportion	Cumulative
Factor1	1.49	0.30	0.30
Factor2	1.00	0.20	0.50
Factor3	0.96	0.19	0.69
Factor4	0.79	0.16	0.85
Factor5	0.76	0.15	1.00

Panel B

Variable	Factor1	Factor2	Unexplained
Finlitir	0.65	-0.22	0.53
Finlitinfl	0.69	-0.01	0.53
Finlitbonds	0.30	0.46	0.70
Finlitrisk	0.66	-0.33	0.46
Finlitmort	0.28	0.80	0.28

4.D. Appendix: Creation of the Inconsistency variables

Preferred basket is the answer to the question in which we ask the respondents the preference between the exclusionary screened and the conventional portfolio. WTP represent whether or not the respondent was willing to pay for screening in his or her portfolio (see Table 4.3). % in stocks is the answer to the question in which we asked the respondents if they would prefer to invest a different percentage in stocks given the basket of stocks is

screened using exclusionary screens compared to their ideal percentage invested in stocks. Inconsistency1 to 3 are the variables we created using the answers those questions. Inconsistency4 takes on the value of one if the respondent makes one of the inconsistent financial choices.

We distinguish three types of beneficiaries, one whose utility function is not influenced by social values, one who gets positive utility from the exclusionary screens proposed, and one who gets negative utility from the proposed screens. All types get positive utility from returns and negative utility from risk. In panel A we take risk aversion into account, panel B follows from panel A. See Table 4.D.1.

Table 4.D.1

Panel A: Answers consistent with three types of rational agents

Utility from screening	Preffered portfolio	WTP	% invested in stocks
No relation	No preference	No	The same
Positive	Screened	No/Neutral/Yes	More/The same
Negative	Conventional	No	Less/The same

Panel B: Definition of ERROR variables

% stocks	WTP	**Inconsistency1**	% stocks	Preference	**Inconsistency2**
	Yes	1		Conventional	0
Less	No	0	Less	Screened	1
	Neutral	1		No preference	1
	Yes	0		Conventional	0
Same	No	0	Same	Screened	0
	Neutral	0		No preference	0
	Yes	0		Conventional	1
More	No	0	More	Screened	0
	Neutral	0		No preference	1

WTP	Preference	**Inconsistency3**	**Inconsistency4**
	Conventional	1	
Yes	Screened	0	1 if $\sum$ERRORi > 0
	No preference	1	
	Conventional	0	
No	Screened	0	
	No preference	0	
	Conventional	1	
Neutral	Screened	0	
	No preference	1	

Chapter 5

5. Values and investments: Evidence from institutional trading responses to news components.

5.1. Introduction

Over the last decades many large institutional investors committed to investing responsibly. A famous initiative is the United Nations-backed Principles for Responsible Investment (UNPRI) for which more than 850 institutions worldwide with assets under management of more than $25 trillion signed up. In the US, about 11% of the stock market investments are managed with an explicit mandate to be responsible according to the Socially Responsible Investment Forum (SIF).[65] There are two common strategies socially conscious investors follow and often combine. First, they overweigh companies with high environmental, social, and governance (ESG) standards and avoid companies with low ESG standards. Second, they exclude companies in controversial or "sinful" industries, such as alcohol, tobacco, or gambling. Institutional investors, who follow an ESG strategy claim that they believe that companies with high ESG standards will outperform in the long-term. A good example is CalPERS, which states in its Sustainability Report (2013, p. 5) "We are long-term investors, so its important that our portfolio companies look out on the horizon as well and implement the kind of ESG policies that will help sustain their future". Similarly, investors who exclude sin or controversial stocks often refer to the long-term risk associated with these industries and products. Based on these statements, it seems that SR-investors are

[65] For more information on UNPRI see http://www.unpri.org/ and for SIF see http://www.ussif.org/

particularly concerned about the long-term, fundamental performance of the companies they invest in.

Another, potentially interrelated motive for holding a responsible portfolio are religious and political values. Hong and Kostovetsky (2012) show that strongly democratic-oriented U.S. investment managers hold less stocks in their portfolio, which are deemed socially irresponsibly than Republicans or investors, who have no strong political views. Kumar and Page (2011) find that norm-constrained institutions only buy sin firms if they expect to earn a high positive abnormal return. The relevance of non-pecuniary motives is supported by the fact that socially responsible portfolios do generally not outperform conventional ones. ESG-strategies have not earned positive abnormal returns in recent years (Borgers, Derwall, Koedijk and ter Horst (2013)). Sin stocks have outperformed comparable stocks (Hong and Kacperczyk (2009)) leading to lower total returns for portfolios that exclude these industries. Just like the pecuniary motives, however, these non-pecuniary motives might be related to a more fundamentally-oriented, long-term investment style. As Nevins, Bearden and Money (2007) show in an experimental setting, people with stronger ethical values are more long-term oriented.

By looking at institutional investors' response to discount rate and cash-flow news, this paper investigates whether socially conscious investors focus more on long-term fundamental information than conventional investors. Cash-flow news, because of its link to production, is relatively more related to firm long-term fundamentals than discount rate news, which can reflect time-varying risk aversion or investor sentiment. To disentangle discount rate and cash-flow news, we apply the return decomposition originally introduced by Campbell (1991) and applied to stock returns by Vuolteenaho (2002). Similar to Cohen, Gompers and Vuolteenaho (2002) we use a vector autoregression (VAR) approach to investigate institutional investors' response to the two different types of news. For the

purpose of our research question, we focus on comparing the response functions of socially conscious and conventional investors. We identify socially conscious investors based on their holdings using three different measures.

First, we use the ESG concern indicators provided by MSCI and calculate a concern score for each portfolio. Portfolios with a low ESG concern score are classified as responsible portfolios, and the ones with a high concern score are classified as conventional portfolios. The second and third classifications are based on the avoidance of sin and controversial stocks. We classify portfolios, which have no sin (controversial) stocks as responsible, and portfolios with large holdings of sin (controversial) stocks as conventional portfolios.

We find that socially conscious investors react significantly to cash-flow news, but not to discount rate news. A 1$ cash-flow shock leads to a trading response of socially conscious investors of 22.3 percent. The reaction to expected return news is statistically insignificant and economically small. Our findings are very different for conventional investors. Conventional investors react much less to cash-flow news than socially conscious investors: a 1$ cash-flow shock is only associated with a trading response of 6.0 percent. They do respond however very strongly to discount rate news. An unexpected return shock of 1$ is associated with a trading response of 33.2 percent. These results are consistent across the three different classifications we use to identify socially conscious investors, and they seem to be a unique characteristic of socially conscious investors. Overall, we can therefore conclude that SR-investors are indeed more focused on fundamental, firm-specific news than conventional investors.

The remainder of this paper is structured as follows: in Section 2 we explain the method we use to pull out news components from stock returns as well as the data we use for the application of this method. Section 3 provides the main results as well as robustness checks and some additional analyses. Finally Section 4 concludes.

5.2. Methodology and data

5.2.1. Return decomposition framework and estimation

Our methodology closely corresponds with the return decomposition framework described in Campbell (1991), Vuolteenaho (2002), and Cohen, Gompers and Vuolteenaho (2002), and others. We follow Vuolteenaho (2002) in that we start with a decomposition of unexpected changes in current individual stock returns into expected return news and cash-flow news. Specifically, we describe unexpected individual stock returns as:

$$r_t - E_{t-1}r_t = \Delta E_t \sum_{j=0}^{\infty} \rho^j e_{t+j} + \kappa_t - \Delta E_t \sum_{j=1}^{\infty} \rho^j r_{t+j} \quad (5.1)$$

Where ΔE_t denotes the change in expectation from t-1 to t, e_t denotes the clean-surplus log accounting return (ROE), r_t the log stock return, κ is an approximation error, and ρ is a constant set to 0.97 (see Vuolteenaho 2002).[66] Since the unexpected return is decomposed into an expected return and a cash-flow component we identify the first component as cash-flow news (N_{cf}) and the second component as expected-return news (N_r).[67] Both components are given by:

$$N_{cf,t} = \Delta E_t \sum_{j=0}^{\infty} \rho^j e_{t+j} + \kappa_t \quad (5.2)$$

$$N_{r,t} = \Delta E_t \sum_{j=1}^{\infty} \rho^j r_{t+j} \quad (5.3)$$

To take the above decomposition to the data we have to make some working assumptions on how investors form expectations about future returns and cash-flows. We follow Vuolteenaho

[66] Specifically, Vuolteenaho(2002) uses a value of 0.967 and shows that his results can be reproduced with values ranging from 0.95 to 1.
[67] For details on the derivation see Cohen et al. (2002).

(2002) and Cohen et al. (2002) among others and use a Vector Autoregression (VAR) to model the forming of these expectations.

We modify the VAR used in Cohen et al. (2002) by separating the institutional ownership component into two parts in order to distinguish between socially conscious and conventional investor ownership. Our VAR model is given by:

$$z_{i,t} = \Gamma z_{i,t-1} + \mu_i + e_{i,t}$$

$$= \Gamma z_{i,t-1} + \mu_{i,t} \tag{5.4}$$

where z denotes a 5x1 vector containing the firm specific state variables. These are given by the market-adjusted log stock return, the market-adjusted log book-to-market ratio, the market-adjusted log accounting return on equity, the market-adjusted fraction of socially conscious investor ownership, and the market-adjusted fraction of conventional investor ownership. Γ denotes a 5x5 matrix of parameters to estimate, and u denotes a 5x1 vector of error terms with covariance matrix Σ. We relax the assumptions of constant parameters and covariance matrix in some of our specifications and allow for some forms of cross-sectional variation.

Using the VAR specification allows us to decompose the unexpected stock return into both news components. The VAR setup implies that:

$$N_{r,it} = e_1' \rho \Gamma (I - \rho \Gamma)^{-1} u_{i,t} \equiv \lambda' u_{i,t} \tag{5.5}$$

where e_1 is a 5x1 vector with one in the first entry and zeros everywhere else and I denotes the identity matrix. We can compute the cash-flow news by:

$$N_{cf,it} = (e_1' + \lambda') u_{it} \tag{5.6}$$

since the total unexpected stock return given by $e_1' u$ equals the sum of both news components we are especially interested in the reactions of conventional and socially conscious investors

to both news components. The reactions are found by regressing the investor ownership fractions on the news components with a constant. Using the VAR structure it is straightforward to get implied regression coefficients.

For the estimation of the VAR we follow the procedure used in Vuolteenaho (2002) and Cohen et al. (2002) and estimate the VAR equation by equation using ordinary least squares (OLS) or weighted least squares (WLS), where we weight each cross section equally. Both procedures give very similar results, and we report the WLS results below. Standard errors are computed via the bootstrap method proposed in Kapetanios (2008).

As in Cohen et al. (2002), the data needed for model estimation comes from the CRSP/Compustat intersection extended with firm-level ownership data taken from CDA/Thomson Reuters institutional ownership holdings. For a thorough description of the holdings- as well as the financial and accounting data, we refer to Appendix 5.A.

The distinction between our VAR approach and that of Cohen et al. (2002) lies in the breakdown of institutional ownership in the VAR. For the execution of our VAR, we must be able determine how much of a firm is owned by, respectively, "sri" and "conventional" institutions, which is discussed in the next section.

5.2.2. Profiling institutional shareholders based on social investment criteria

To understand whether socially conscious institutions respond differently to news than do conventional institutions, we must first profile each institutional owner along relevant social dimensions in investing. To accomplish this goal we use ownership data from the 13F SEC filings of institutional investors with more than $100 million of discretionary assets under management and discloses for each institutional money manager (i) the amount of

shares held of each firm (j)[68]. In order to derive a social profile for each institutional manager, we match their holdings in the second quarter of each year with the annual firm- level ESG ("Environmental, Social, and Governance") indicators from the Morgan Stanley Capital Indexes STATS (MSCI STATS) database. STATS provides these indicators on a yearly basis, starting in 1991 with an analyses of S&P500 stocks and constituents of the Domini 400 social index.

Given that literature and anecdotal evidence suggest that socially conscious institutions have a history of avoiding social controversies, we collect for our main analysis the "concerns" indicators from all of the ESG categories covered in MSCI STATS (community, diversity, employee relations, environment, human rights, corporate governance, and product)[69]. Because the STATS universe increases over time, we only use the indicators for S&P500 stocks, keeping our measurement procedure stable over time. With the holdings data and the ESG concerns data that are available, we value-weight the concerns for each institutional portfolio:

$$IOconcerns_{i,t} = \sum_{j=1}^{J} \frac{(S\&P500)shares_{j,t} * price_{j,t} * Concerns_{j,t}}{total\ (S\&P500)\ assets\ under\ management_{i,t}} \quad (5.7a)$$

where *Concerns* is the size-adjusted sum of all ESG concerns indicators.[70] We use institutions that have invested in at least 100 firms in the S&P500[71]. This results in a sample of 1476 institutions and 9681 institution-years. For each institution, we calculate the time series mean

[68] The holdings data have some recording errors, which lead to firms with institutional ownership greater than 100 percent. We deal with this issue in two ways. At the institutional level we drop observations for which an institution claims to hold more than 110 percent of shares outstanding (due to rounding issues we cannot delete institutions that claim to hold more than 100 percent). And at the firm level we follow Lewellen (2001) and trim ownership observations at the 100 percent level so that each firm can only be owned for 100 percent by institutional owners and change the holdings for the sri and conventional owners accordingly.

[69] See 6. General Appendix for a description of the MSCI STATS database.

[70] We apply the size adjustment to account for the positive correlation between size and the number of social responsibility strength- and concern indicators. We size-adjust by subtracting the mean of each environmental and social score within each size decile (based on market cap) in each year.

[71] Using only institutions that have at least invested in 0 or 25 MSCI STATS rated firms does not qualitatively alter the results.

of *IOconcerns* and define institutions in the bottom 1/3rd of the distribution of the mean-adjusted score, hence with relatively few investments in firms with ESG concerns, as socially conscious (sri) institutions.[72] All other institutions are classified as conventional institutions.

In additional analyses, we verify to what extent our results hold if we use different definitions of controversial assets to identify socially conscious institutions. Next to ESG indicators, MSCI STATS maintains controversial business indicators (tobacco, alcohol, gambling, weapons and military, and nuclear operations) that we can use to identify "sin stocks", which according to studies are likely to be shunned by U.S. institutions that are sensitive to social norms (e.g. Hong and Kacperczyk 2009). Also SRI mutual funds have a history of avoiding these stocks (e.g. Hong & Kostovetsky 2012, and Borgers et al. 2014). Following Borgers et al. (2014), we adopt two definitions of "sin stocks". Next to the classical working definition that includes stocks from tobacco, alcohol and gambling sectors (see, for example, Hong and Kacperczyk 2009), we broaden the sin stock universe to include all possible controversial business indicators from MSCI STATS (tobacco, alcohol, gambling, weapons/military, and nuclear operations). We define the broader set as "controversial stocks".[73] The definitions are formalized as follows:

$$IOsin_{i,t} = \sum_{j=1}^{J} \frac{(S\&P500)shares_{j,t} * price_{j,t} * Dsin_j}{total\ (S\&P500)\ assets\ under\ management_{i,t}} \quad (5.7b)$$

$$IOcontrov_{i,t} = \sum_{j=1}^{J} \frac{(S\&P500)shares_{j,t} * price_{j,t} * Dcontroversial_j}{total\ (S\&P500)\ assets\ under\ management_{i,t}} \quad (5.7c)$$

where *Dsin* is one for sin stocks (alcohol, gambling, tobacco) and *Dcontroversial* is an indicator for controversial stocks (sin, weapons, military and defense, and nuclear operations).

[72] Using the raw *IOconcerns* or [time series mean / time series standard deviation] of *IOconcerns* does not alter our findings significantly.

[73] Among these additional stocks, especially stocks of weapons manufacturers are increasingly being blacklisted by institutional investors domiciled in the U.K. and continental Europe.

Using *IOsin (IOcontrov)* we define socially conscious institutions as those in the bottom (top) 1/3rd of the distribution.

By classifying each institution as either socially conscious (SRI) or conventional, we are able to determine annually, for each firm in the sample, the percentage of total equity that is held by each of these institutional shareholder categories. As these classifications play a crucial role in our analyses we provide a deeper look into their scores in Section 4.

5.3. Empirical results

5.3.1. Summary statistics

Panel A of Table 5.1 shows descriptive statistics on relevant firm-level variables. For each firm, we calculated annually the log of one plus the firm's stock return (ret), the natural logarithm of the firm's book-to-market ratio (b2m), and the natural logarithm of one plus the firm's accounting return on equity (roe). In addition, we computed the fraction of a firm that is held by institutions labeled socially conscious (SRI) and those labeled conventional (conv), using *IOConcerns* to distinguish institutions.

On average SRI institutions hold 11.2 percent of the equity capital of each firm, while 21.7 percent is held by conventional institutions this difference is by construction as we define SRI institutions as the all institutions in the bottom 1/3rd of *IOconcerns* and conventional institutions are all others. On average, the firms in our sample produced a negative annual stock return of 2.4 percent. The distribution of returns is somewhat left skewed as the median firm returned positive 1.8 percent. The same holds for the return on equity with a mean and median of -1.8 percent and 8.2 percent respectively. For the sake of completeness, Panels B and C show correlations between the variables – on a market-adjusted basis - as in Cohen et al. (2002).

Table 5.1 Descriptive statistics

This table presents descriptive statistics of the data we use in the VAR model. The data come from the intersection of the CRSP and Compustat datasets matched to ownership data from CDA Thomson Reuters that was on its turn matched to the MSCI STATS database. In Panel A we present statistics on the natural logarithm of excess returns (ret), the natural logarithm of the book-to-market ratio (b2m), the natural logarithm of the GAAP accounting return on equity (roe) and the fractions of the firms owned by socially responsible (sri) and conventional (conv) investors. Panel B presents contemporaneous correlations of market adjusted data (market adjustment is indicated by an underline). Panel C presents First order cross- and autocorreltaions.

Panel A: Basic descriptive statistics

	Mean	Std. Dev.	Min.	25%-pct	Median	75%-pct	Max.
ret	-0.024	0.452	-2.736	-0.231	0.018	0.221	3.084
b2m	-0.495	0.654	-2.215	-0.912	-0.501	-0.103	4.147
roe	-0.018	0.433	-2.302	-0.005	0.082	0.14	4.223
sri	0.112	0.107	0.000	0.021	0.083	0.174	0.863
conv	0.217	0.158	0.000	0.081	0.192	0.332	0.973

Panel B: Contemporaneous correlation, market adjusted data

	ret	b2m	roe	sri	conv
ret	1.000	-0.504	0.140	0.082	0.030
b2m	-0.504	1.000	-0.212	-0.155	-0.056
roe	0.140	-0.212	1.000	0.120	0.102
sri	0.082	-0.155	0.120	1.000	0.161
conv	0.030	-0.056	0.102	0.161	1.000

Panel C: First-order cross- and autocorrelations, market adjusted data

	ret	b2m	roe	sri	conv
ret	-0.099	0.181	-0.073	-0.093	-0.072
b2m	-0.157	0.285	-0.063	-0.006	0.039
roe	0.206	-0.34	0.161	0.072	0.026
sri	0.097	-0.127	0.063	0.306	0.091
conv	0.048	-0.062	0.041	0.119	0.359

5.3.2. VAR outcome and response coefficients

Table 5.2 shows the VAR coefficients, which we obtained by estimating specification (5.4). Consistent with earlier studies, the VAR coefficients in the first row of Table 5.2 indicate a positive association between stock return and, respectively, past annual return, the book-to-market ratio, accounting return on equity, the percentage of stocks held institutions (both SRI and conventional). As for the ownership variables, the last columns in Table 5.2 indicate that

both ownership by socially conscious institutions and ownership by conventional institutions is significantly greater in firms with higher past 12-month return, higher accounting return on equity, and higher level of past SRI and conventional institutional ownership. Unlike conventional institutional ownership, ownership by socially conscious institutions is lower when the book-to-market ratio is higher. This differential sensitivity is in line with the stylized fact that socially conscious investors tend to experience tilts to "growth" or "glamour" stocks, possibly because they avoid stocks with typical value characteristics (e.g. Hong and Kacperczyk 2009).

Table 5.2 VAR coefficients

The table reports the VAR parameters estimated form the annual panel. The model state variables include the market-adjusted log stock return, r (the first element of the state vector z); market-adjusted log book-to-market ratio, b2m (the second element); market-adjusted log profitability, roe (the third element); and market-adjusted fraction of shares outstanding owned by respectively socially conscious institutions, sri (the fourth element) and conventional institutions conv. The parameters correspond to the following system:

$z_{i,t} = \Gamma z_{i,t-1} + \mu_{i,t}, \Sigma = E(u_i u_{i,t}')$

We report two numbers for each parameter. The first number is a weighted least squares estimate of the parameter, where observations are weighted such that each cross-section receives an equal weight. The second number (in parentheses) is a robust standard error computed using Rogers's (1983, 1993) method (details of the method are described by Vuolteenaho, 2002).

VAR coefficients										
	Γ					Σ				
ret	0.163	0.127	0.098	0.055	0.251	0.181	-0.118	0.028	0.005	0.004
	(0.016)	(0.009)	(0.014)	(0.067)	(0.049)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)
b2m	0.078	0.794	0.032	0.056	-0.093	-0.118	0.169	0.009	-0.004	-0.003
	(0.017)	(0.011)	(0.020)	(0.057)	(0.039)	(0.002)	(0.003)	(0.001)	(0.000)	(0.000)
roe	0.183	0.079	0.558	0.326	0.226	0.028	0.009	0.110	0.001	0.002
	(0.016)	(0.010)	(0.015)	(0.076)	(0.051)	(0.001)	(0.001)	(0.004)	(0.000)	(0.000)
sri	0.012	-0.004	0.004	0.746	0.088	0.005	-0.004	0.001	0.003	0.000
	(0.002)	(0.001)	(0.001)	(0.011)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
conv	0.012	-0.001	0.005	0.100	0.884	0.004	-0.003	0.002	0.000	0.004
	(0.002)	(0.001)	(0.001)	(0.010)	(0.007)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

In Table 5.3, Panel A, we show the covariance (correlations in bottom left) between the news variables and the two institutional ownership variables that are central to our analysis. Consistent with Cohen et al. (2002), firm-level market-adjusted return is more driven by cash-flow news than by expected-return news, and the news variables are positively correlated. The variances of expected-return news (0.176) and cash-flow news (0.393) are both larger than those that Cohen et al. (2002) report based on their sample, but the positive correlation between the news variables is similar (0.738).

Table 5.3 Responses of prices and institutional ownership to news

The table reports derived statistics calculated from the VAR specification of Table 5.2. The VAR specification has the structure: $z_{i,t} = \Gamma z_{i,t-1} + \mu_{i,t}$, $\Sigma = E(u_i u_{i,t}')$

The model variables include the market-adjusted log stock return, market-adjusted log book-to-market ratio, market-adjusted log accounting return on equity, and market-adjusted institutional-ownership fractions from "socially conscious" and "conventional" investors respectively. Panel A reports the covariance and correlation matrices of expected-return news, cash-flow news, and institutional-ownership shock. The upper-left section (including diagonal) of the panel shows covariances and the lower-left section correlations. Panel B shows regression of return and institutional-ownership shock on cash-flow and expected-return news.

Panel A: Covariances and correlations of news and institutional ownership				
	ern	cfn	sri	conv
ern	0.176	0.194	0.007	0.015
	(0.030)	(0.035)	(0.001)	(0.002)
cfn	0.738	0.393	0.012	0.019
	(0.033)	(0.042)	(0.001)	(0.002)
sri	0.537	0.360	0.003	
	(0.040)	(0.025)	(0.000)	
conv	0.744	0.498		0.004
	(0.050)	(0.037)		(0.000)
Panel B: Regressions				
			cfn	ern
u_r			0.506	
			(0.037)	
u_{SRI}			0.031	
			(0.002)	
u_{CONV}			0.048	
			(0.003)	
u_{SRI}			0.025	0.011
			(0.002)	(0.006)
u_{CONV}			0.013	0.072
			(0.004)	(0.011)

We now turn our attention to responses of return and institutional ownership to cash-flow news. In the first row of Table 5.3, Panel B, we first present the estimated response of returns to cash-flow news. The coefficient of 0.506 indicates that stock prices move less than one-to-one with a $1 cash-flow shock. Our estimate of a $0.494 (1-0.506) underreaction is somewhat stronger than the $0.41 underreaction that Cohen et al. (2002) report based on a different sample period.

The subsequent coefficients indicate that ownership by socially conscious institutions and ownership by conventional institutions relate positively to cash-flow news. Given the underreaction of return to news, these results suggest that institutions buy on positive cash-flow news in order to exploit the apparent underreaction. At first glance, they seem to do so to a different degree. The estimate response of *sri* to cash flow news equals 0.031, whereas the coefficient on *conv* is 0.048. However, we hasten to comment that the average percentage held by conventional institutions is by construction roughly twice as large as the percentage held by socially conscious institutions (as indicated in Table 5.1). Once the coefficients are scaled by mean ownership levels, *sri* (*conv*) increase their positions by 27.7 percent (22.1 percent) on average after a cash-flow news shock of 1$. Given our model specifications and that we scale the coefficients simply by mean ownership over firms and time the responses to cash-flow news in this univariate setting are not extremely different.

Although the abovementioned positive coefficients for the ownership variables could be taken to imply that both types of institutions buy stocks on positive cash-flow news in order to profit from underreaction, they could also point to a mere preference for holding stocks after positive cash-flow news (Cohen, Gompers and Vuolteenaho 2002). To shed more light on the nature of the response coefficients, we report in the last lines of Table 5.3 the ownership regressions that include both cash-flow news and expected-return news.

In the case of ownership by conventional institutions, the magnitude of the coefficient on cash-flow news decreases to 0.013 when expected-return news is added as a regressor, and ownership loads significantly positively on expected-return news. That the coefficient on expected-return news is economically and statically significant suggests that conventional institutions buy on positive cash-flow news due to the positive expected return implied by the under reaction. Implicit in this interpretation is the assumption that positive expected return-news captures mispricing (under reaction) because it is derived from market-adjusted data.

In the case of ownership by socially conscious institutions, however, we find that expected-return news hardly subsumes the positive response to cash-flow news, and ownership does not significantly react to expected-return news. In other words, socially conscious institutions appear to be sensitive to cash-flow news, but not because of the higher expected return associated with under reaction to cash-flow news.

To formally test whether the effects described above are different we need to scale the response coefficients in order to compare them. Therefore, we scale the response coefficients of *sri* and *conv* from the bivariate regression results presented in Table 5.3 by the respective mean ownership levels. For each bootstrap repetition we compute the difference between the mean-scaled responses to both news components. This gives us the mean scaled differences as well as their standard deviations. Socially conscious institutions react 15.2 percent (st.dev. 1.0 percent) stronger to cash-flow news while they react 17.0 percent (st.dev. 6.9 percent) less to expected return news. With *t*-statistics of 15.2 and 2.5 respectively both the differences in response are statistically significant at conventional levels.

The observed differential response to the news variables can be seen as an indication that socially conscious institutions are, compared to conventional institutions, mainly focused on the permanent components of returns because of a more long-term orientation. One would expect that long-term investors are mainly concerned about permanent shocks to their wealth,

as manifested in cash-flow news (see, e.g., Campbell, Polk and Voulteenaho 2010). They would arguably respond less to expected-return news because of mean reversion in returns that are unrelated to fundamentals. This interpretation concerning investors' response to the news variables can be further understood by means of Figure 5.1, which shows what happens to cumulative (market-adjusted) returns if we shock the VAR with a 25 percent return without any cash-flow news, and what happens if we introduce a shock that is 25 percent cash-flow news. Over time the cash-flow news shock gets priced in as the cumulative market-adjusted) return goes to 25 percent. However, in the absence of cash-flow news, the cumulative market-adjusted return is close to 0 percent.[74]

5.3.3. Alternative definitions of social controversies.

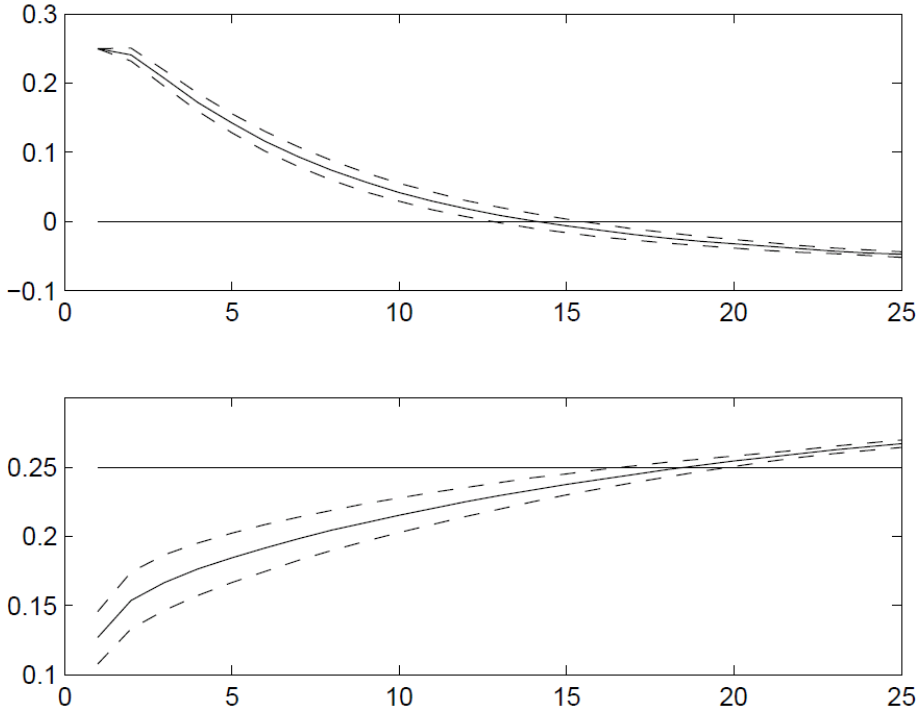
We now verify whether the classification of institutions as either socially conscious (conventional) based on alternative definitions of social controversies affects our main results. To do so, we recompute *sri* and *conv* based on the alternative definitions described in Section 2.2, and estimate their responses to cash-flow news and expected-return news. The results in the first set of columns that are presented in Table 5.4 emerge from using *IO*sin to distinguish socially conscious institutions from conventional institutions. The second set of columns represents the outcome of using *IOcontrov* to make this distinction. Socially conscious institutions are those that comprise the bottom 1/3rd of the distribution of these alternative institutional investor scores, and conventional institutions are those that below to the remaining 2/3rd. To facilitate a comparison with our baseline results, we summarize the relevant ownership response coefficients from Table 5.3 in the third set of columns that are presented in Table 5.4.

[74] These patterns are very similar to those of Voulteenaho (2002) and those of Cohen et al. (2002). The reported price response over time to a 25-percent shock in return is consistent with an initial momentum effect a la Jegadeesh and Titman (1993) and subsequent reversal in the spirit of De Bondt and Thaler (1985).

Figure 5.1 Cumulative return responses to 25 percent return shocks

The figure contains two impulse-response functions computed from the VAR system of Table 2. The figures plot the cumulative returns over years if we shock the VAR model with a 25 percent return in the *absence* of any cash-flow news (Top) or give an impulse of 25 percent cash-flow news (bottom).

The 25 percent return shock is generated by setting the first element of the VAR-error vector to 0.25. The other elements of the VAR-error vector are set to their conditional expectations, conditional on the first element being equal to 0.25 and cash-flow news equal to zero. The typical 25 percent cash-flow news is induced by setting the VAR-error vector to a constrained maximum likelihood value, imposing the constraint that cash-flow news equals 0.25. Dashed lines denote –/+2 standard-error bounds.



Under these alternative definitions, the percentage of equity that is held by socially conscious institutions is smaller than in our baseline case. Using sin (controversial) stock investments to identify socially conscious investments yields average ownership of such investors of 5.7 percent (4.7 percent) in the average firm compared to 27.1 percent (28.2 percent) for conventional investors.

We now turn to the response coefficients to cash-flow news and expected-return news for both types of institutional ownership. Because ownership by socially conscious

institutions is smaller when derived from either *IOsin* or *IOcontrov*, compared to that derived from *IOConcerns*, the response coefficients are in magnitude smaller than those reported in our main analysis. Nevertheless, the main inferences remain unchanged: socially conscious institutions buy on positive cash-flow news and do not respond to expected-return news, whereas conventional institutions respond positively to both cash-flow news and expected-return news.

In another robustness check we allow institutions to switch groups (from socially conscious to conventional and the other way around) by classifying institutions based on their raw *IOscore* instead of the time series mean of the score. In Table 5.5 we report the results and show that the results are not sensitive to this modification.

Table 5.4 Different identification of SR investors: Sin and Controverial investments

In this table we use three different definitions of the IOscore to identify socially conscious investors. In the first two rows we present the average ownership by "sri" and "conv" investors identified according to one of three definitions (less investments in respectively Sin stocks, Controversial stocks, or stocks with more ESG concerns). The rows below present results from bivariate regressions as presented in Table 5.3. Standard errors derived from 1000 bootstrap replications are given in parentheses below the implied VAR coefficients.

	SIN		CONTROV		CONCERNS	
	sri	conv	sri	conv	sri	conv
Mean	0.057	0.271	0.047	0.282	0.112	0.217
Std	0.060	0.210	0.064	0.200	0.107	0.158
cfn	0.006	0.031	0.009	0.033	0.025	0.013
	(0.001)	(0.006)	(0.001)	(0.005)	(0.002)	(0.004)
ern	0.005	0.076	0.007	0.070	0.011	0.072
	(0.004)	(0.013)	(0.005)	(0.012)	(0.006)	(0.011)

Table 5.5 Different identification of SR investors: Raw IOscore

In this table we use the raw IOscores instead of the timeseries mean IOscore for each institution. We again present three different definitions of the IOscore to identify socially conscious investors. In the first two rows we present the average ownership by "sri" and "conv" investors identified according to one of three definitions (less investments in respectively Sin stocks, Controversial stocks, or stocks with more ESG concerns). The rows below present results from bivariate regressions as presented in Table 5.3. Standard errors derived from 1000 bootstrap replications are given in parentheses below the implied VAR coefficients.

	SIN		CONTROV		CONCERNS	
	sri	conv	sri	conv	sri	conv
Mean	0.059	0.269	0.099	0.229	0.095	0.233
Std	0.064	0.205	0.093	0.176	0.096	0.169
cfn	0.011	0.029	0.017	0.022	0.022	0.018
	(0.001)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)
ern	0.008	0.073	0.012	0.070	0.012	0.069
	(0.005)	(0.013)	(0.006)	(0.012)	(0.007)	(0.012)

5.4. Validation of IOscores

5.4.1. IOscores and political values

Using controversial assets to identify socially conscious and conventional investors is central to our analysis, but a concern is that our measure might not capture the explicit preference of investors' avoidance for non-controversial assets. Therefore we provide a simple exercise that relates our measure to a proxy for political preferences, which have been shown to be related to preferences for social investments.[75] Hong and Kostovetsky (2012) use political contributions of mutual fund managers as a measure of political preferences and show that (strong) Democrat contributors are less likely to invest in Controversial stocks (using a similar definition as ours) and stocks with Employee Relations- or Workforce

[75] For a thorough analysis of the relation between social and political preferences see Hong and Kostovetsky (2010).

Diversity concerns. Furthermore, Di Giuli and Kostovetsky (2013) argue that location of a firm influences the investment decisions since most outside stakeholders are located in proximity to the firm. They show that firms located in more Democrat oriented states score higher on measures of corporate social responsibility. Since we don't have manager-level data, we follow Di Giuli and Kostovetsky and use the location as a proxy for political preferences. We search for the name and location of institutions in the top and bottom of the IOmeasures, when no specific division, and therefore location, is indicated in the data we use headquarters location. Subsequently we add election result data for all presidential elections that took place in during our sample period; 1992, 1996, 2000, 2004, and 2008. For each state we add up the number of times Democrats won the elections (e.g. California = 5, Florida = 2, Utah = 0) and appoint this score to the institutions as shown for the 10 Top and Bottom institutions using the IOconcerns measure in Table 5.6.

If our measure picks up the correlation between social and political preferences reported in the literature, we should see that institutions with higher exposures to stocks with ESG concerns are located in states with a more Republican preference.

From Table 5.7, Panel A, it becomes clear that the top of our SRI group is located in more democrat-oriented states whereas the institutions with the highest exposure to assets with ESG concerns are headquartered more in states with a Republican voting majority. This effect is statistically significant at the 1 percent and 5 percent level for the controversial- and sin stock measures respectively. Moreover we confirm these findings in Table 5.7 Panel B where we compare all institutions that are in the lowest 1/3rd of the distribution to the institutions in the highest 1/3rd *for each* IOmeasure. This exercise shows that our measures to identify social values in institutional investment portfolios are correlated with political preferences albeit a crude measure.

Table 5.6: Top/Bottom IOconcerns institutions and state level election results

This table reports a list of 10 institutions with the lowest and highest IOconcerns measure. For each institution we hand collect location data (if a specific division location is not indicated in the name we use the headquarters location). For each state we then provide the number of times the democrats won the presidential elections, ElectionWinDem.

	Manager name	STATE/Location	ElectionWinDem
Panel A: High exposure to STATS concerns			
4.763	EVERCORE TRUST COMPANY, N.A.	New York	5
4.756	ALBION MANAGEMENT GROUP	Connecticut	5
3.112	PINNACLE MANAGEMENT & TRUST CO	Texas	0
3.067	WEBSTER TRUST COMPANY NA	Arizona	1
2.854	TD CAPITAL MANAGEMENT LLC	Tennessee	2
2.479	MACQUARIE INVESTMENT MGMT LTD.	Australia	na
2.428	MECHANICS BANK TRUST DEPT	California	5
2.414	FIDUCIARY GROUP	Europe (Gibraltar)	na
2.266	MANUFACTURERS NATL CORP	Texas	0
2.218	CAMBRIDGE INV RES ADVISORS INC	Iowa	4
TotalWinDem			22
MeanWinDem			2.750
Panel B: Low exposure to STATS concerns			
-2.657	RBC TRUST CO (INT'L) LTD	Canada	na
-1.966	FCM INVESTMENTS, LP	Texas	0
-1.832	ENGEMANN ASSET MANAGEMENT	California	5
-1.541	SANTA BARBARA ASSET MGMT, LLC	California	5
-1.455	CITIZENS ADVISERS INC	New Hampshire	4
-1.391	CENTURION ALLIANCE, INC	California	5
-1.375	VOYAGEUR ASSET MANAGEMENT INC.	Minnesota	5
-1.332	TRUST CO BANK OF GEORGIA	Georgia	1
-1.318	PADCO ADVR II, INC	Maryland	5
-1.296	SPIRIT OF AMERICA MGMT CORP	New York	5
TotalWinDem			35
MeanWinDem			3.889

Table 5.7: Top/Bottom IOmeasures and state level election results: difference tests

Panel A of this table reports on difference tests of the measure MeanWinDem as explained in Table 5.1. In panel B we compare MeanWinDem for all institutions that are in the Highest 1/3rd of the distribution for each IOmeasure (Concerns, Sin, and Controversial) to all institutions that are in the Lowest 1/3rd. The P-values between squared brackets are based on 1 tailed t-tests.

Panel A: MeanWinDem per IOmeasure		
Ioconcerns	High exposure	2.750
	Low exposure	3.889
	Difference	-1.139
	P-value	[0.141]
Iosin	High exposure	3.250
	Low exposure	4.900
	Difference	-1.650
	P-value	[0.027]
Iocontroversial	High exposure	1.700
	Low exposure	5.000
	Difference	-3.300
	P-value	[0.000]

Panel B: MeanWinDem all IOmeasures		
All measures	Highest 1/3	3.450
All measures	Lowest 1/3	4.050
	Difference	-0.600
	P-value	[0.029]

5.4.2. Institution types

Another concern that we want to address is that we might overweight certain institution types. To address this concern we analyze the relation between our IOmeasures and institution type data from CDA/Thomson Reuters. They assign one of five types to an institution; (1) bank, (2) insurance company, (3) investment company (mutual fund), (4) investment advisor, and (5) other. Types (1)-(3) are self explanatory, the investment advisor category includes most of

the large brokerage firms, and the "other" category includes pension funds and university endowments. Because the data provider indicates that after 1998 institutions can be wrongly classified as type (5) we also use updated classification data provided by Brian Bushee on his personal website[76]. Besides being updated after 1998, this measure also separates the "other" type into private pensions, public pensions, and endowments. All institutions that do not fall into one of the above mentioned groups were classified as "miscellaneous".

We create dummy variables for each of the institution types and run pooled OLS regressions with the three IOscores on the institution type dummies, total equity, total number of stocks, total equity per stock, and year fixed effects. The results are reported in Table 5.8, the banking sector has a higher exposure to ESG concerns and stocks with more controversial stocks than do independent investment advisors (+3 percent higher exposure within the S&P500 investments). This effect is driven by banks investing more in firms with operations in the weapons, military and defense, and nuclear energy sectors. Overall, the institution types do not differ too much in their exposures to controversial stocks or stocks with ESG concerns as only 7 out of 33 institution dummies are significantly different from zero at conventional levels. Moreover, the results differ per measure.

Table 5.8: Explaining IOscores by institution type

This table presents the results from OLS regressions of IOmeasures (Concerns, Sin, and Controversial) on style dummies taken directly from CDA/Thomson Reuters or style dummies from Brian Bushees personal website. As control variables we add the natural logarithm of Total Equity of the institution (manager), the natural logarithm of the Number of Stocks in the portfolio, Average Equity per stock in mln $US, and year fixed effects. Since we use mutually exclusive types we use the most common type, "independent investment advisors", as the base case. T-stats, presented in parentheses, are derived from standard errors clustered at the institution and year level.

[76] http://acct.wharton.upenn.edu/faculty/bushee/IIclass.html

Table 5.8 continued

	IOconcerns		IOsin		IOcontrov	
OrigStyleBank	0.224***		0.002		0.033***	
	(5.081)		(0.904)		(6.653)	
OrigStyleInsurance	-0.007		-0.000		0.002	
	(-0.189)		(-0.286)		(0.391)	
OrigStyleInvestment	0.027		0.001		-0.001	
	(0.602)		(0.842)		(-0.170)	
OrigStyleOther	-0.017		0.002**		0.002	
	(-0.609)		(2.262)		(0.481)	
BusheeBank		0.262***		0.001		0.032***
		(5.945)		(0.241)		(7.452)
BusheeCorpPens		0.141***		-0.000		0.012
		(3.309)		(-0.282)		(1.207)
BusheeInsurance		-0.002		-0.002		0.008
		(-0.038)		(-1.082)		(1.575)
BusheeInvestment		-0.023		0.001		0.007
		(-0.510)		(0.365)		(0.890)
BusheePublicPens		0.127***		-0.000		0.008
		(4.500)		(-0.323)		(1.384)
BusheeUniv&Foundation		0.040		0.001		0.004
		(1.026)		(0.194)		(0.398)
BusheeMiscellaneous		0.037		0.005		-0.024***
		(0.618)		(1.600)		(-3.839)
Observations	9,681	9,410	9,681	9,410	9,681	9,410
R-squared	0.038	0.054	0.024	0.025	0.282	0.292
Controls and constant	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

5.4. Conclusion

Research on social values in financial markets concentrates on characteristics of assets which are over- or underweighted in investment portfolios given a set of norms or values (see, e.g. Hong and Kacperczyk 2009, Hong and Kostovetsky 2012, Kumar and Page 2011). In this paper we argue that besides having tastes against certain types of stocks, social values also influence the investment behavior for the assets that investors do hold. More specifically we argue that investors who are more socially conscious have a different investment orientation in that they are less sensible to non-persistent stock price effects not driven by fundamentals.

Since social values are known to manifest themselves in tastes for or against controversial assets we use the holdings of US institutional investors as means to indentify two investor types. We define *socially conscious* investors as institutions that have few assets under management invested in stocks with Environmental, Social, and Governance controversies and all other institutions as *conventional*. Subsequently we use a vector autoregression (VAR) framework to identify cash-flow (permanent) and expected return (temporary) news components of stock returns (following e.g. Vuolteenaho 2002 and Cohen et al. 2002) and analyse the relative responses of the investor types to both news components.

Our main results indicate that *socially conscious* investors respond 17.0 percent less to expected return news while these investors respond 15.2 percent stronger to cash-flow news. This evidence is in line with the view that social values do not only influence tastes against stocks with social controversies but also the investment decisions for all assets in the portfolio on average.

5.A. Appendix: Data

This appendix is taken from Cohen et al. (2002) and adjusted for the data used in this paper.

5.A.1. CRSP–COMPUSTAT data

The basic data come from the CRSP–COMPUSTAT intersection. The Center for Research in Securities Prices (CRSP) monthly stock file contains monthly prices, shares outstanding, dividends and returns for NYSE, AMEX, and NASDAQ stocks. The COMPUSTAT annual research file contains the relevant accounting information for most publicly traded US stocks. In order to be included in our sample, a firm-year must satisfy the following COMPUSTAT data requirements. First, we require all firms to have a December fiscal-year end of t-1; in order to align accounting variables across firms. Second, a firm must have t-1; t-2; and t-3 book equity available, where t denotes time in years. A number of CRSP data requirements must also be satisfied. A valid market equity figure must be available for t-1; t-2; and t-3: We require that there is a valid trade during the month immediately preceding the period t return. This requirement ensures that the return predictability is not spuriously induced by stale prices or other similar market micro-structure issues. We also require at least one monthly return observation during each of the preceding five years, from t-1 to t-5: In addition, we screen out clear data errors and mismatches by excluding firms with t-1 market equity less than $10 million and book-to-market more than 100 or less than 1/100. We carefully avoid imposing any COMPUSTAT or CRSP requirements on year t data, because these data are used in the dependent variables of our regressions.

The stock returns are calculated as follows. The simple stock return is an annual value-weighted return on a firm's common stock issues (typically one). If no return data are available, we substitute zeros for both returns and dividends. Annual returns are compounded from monthly returns, recorded from the beginning of July to the end of June.

Delisting returns are included when available in CRSP. If a firm is delisted but the delisting return is missing, we investigate the reason for disappearance. If the delisting is performance-related, we assume a -30% delisting return. Otherwise, we assume a zero delisting return. The delisting-return assumptions are based on Shumway's (1997) results. Shumway tracks a sample of firms whose delisting returns are missing from the CRSP data and finds that performance-related delistings are associated with a significant negative return, on average approximately -30%. This assumption is unimportant to our final results, however. Market equity (combined value of all common stock classes outstanding) is taken from CRSP as of the end of June. If the year t market equity is missing, we compound the lagged market equity with return without dividends.

For book equity, we prefer COMPUSTAT data item 60, but if it is unavailable we use item 235. Also, if short- and/or long-term deferred taxes are available (data items 35 and 71), we add them to book equity. If both data items 60 and 235 are unavailable, we proxy book equity by the last period's book equity plus earnings less dividends. If neither earnings nor book equity is available, we assume that the book to market ratio has not changed and compute the book equity proxy from the last period's book-to-market and this period's market equity. We treat negative or zero book equity values as missing.

GAAP ROE is the earnings over the last period's book equity, measured according to the US Generally Accepted Accounting Principles. We use the COMPUSTAT data item 172, earnings available for common. When earnings are missing, earnings is computed as the change in book equity plus dividends. In every case, we do not allow the firm to lose more than its book equity. That is, we define the net income as maximum of reported net income (or clean-surplus net income, if earnings are not reported) and negative of the beginning of the period book equity. Hence, the minimum GAAP ROE is truncated to -100%.

We calculate leverage as book equity over the sum of book equity and book debt. The book debt is the sum of debt in current liabilities (34), total long-term debt (9), and preferred stock (130). The identities necessitate the use of log transforms of stock return, profitability and the book-to-market ratio. The log transformations can cause problems if some stock returns and/or ROEs are close to -100% or if some of the book-to-market ratios are close to zero or infinity. We follow Vuolteenaho (2002) and solve this complication by redefining the firm as a portfolio of 90% common stock and 10% Treasury-bills using market values. Every period, the portfolio is rebalanced to these weights. This affects not only stock return and accounting return on equity, but also the book-to-market equity, pulling this ratio slightly towards one. After adding this risk-free investment, the ratios and returns are sufficiently well behaved for log transformations. Simple market and accounting returns on this portfolio closely approximate simple returns on the firm's common stock only. The accounting identities hold for the transformed quantities. Furthermore, this transformation method is superior to purely statistical transformations (such as the Box-Cox transformation), because the transformed quantities still correspond to an investment strategy.

5.A.2. CDA/Thomson reuters data

A 1978 amendment to the Securities and Exchange Act of 1934 required all institutions with greater than $100 million of securities under discretionary management to report their holdings to the SEC. Holdings are reported quarterly on the SEC's form 13F, where all common-stock positions greater than 10,000 shares or $200,000 must be disclosed. These reports are available in electronic form back to 1980 from CDA/Spectrum, a firm hired by the SEC to process the 13F filings. Our data include the quarterly reports from the second quarter

of 1991 through the second quarter of 2009. Throughout this paper, we use ‘‘institution,’’ as a synonym for ‘‘an institution that files a 13F.’’

On the 13F, each manager must report all securities over which they exercise sole or shared investment discretion. In cases where investment discretion is shared by more than one institution, care is taken to prevent double counting. Spectrum officials have told us that they believe that duplication is rare. Once an institution enters the 13F sample, it is assigned a manager type by Spectrum. The five types are (1) bank, (2) insurance company, (3) investment company (mutual fund), (4) investment advisor, and (5) other. The first three categories are self-explanatory. The investment advisor category includes most of the large brokerage firms, and the “other” category includes pension funds and university endowments. These categorizations are not always precise, though. For example, brokerage firms with mutual fund subsidiaries will fall into category (3) if the mutual funds are deemed by Spectrum to make up more than 50% of the total 13F assets for that manager and into category (4) otherwise. Spectrum does not provide information to allow more precise partitioning of the data. It is also possible for a manager to be reclassified over time if Spectrum determines that the institution’s main business has changed.

The Spectrum 13F holdings file contains three columns: date, CUSIP code, identifier for the institution, and number of shares held in that stock by that institution on that date. All dates are end-of-quarter (March 31, June 30, September 30, or December 31). For each CUSIP and date we simply sum up the shares held by all institutions in the sample to get total institutional holdings of the security at the end of that quarter. We do not drop firms without SPECTRUM data from our sample. Firms with no SPECTRUM data are recorded as having zero institutional ownership. We then match each CUSIP to a CRSP PERMNO, the permanent number CRSP assigns to that security. Holdings associated with CUSIPs for which we found no associated PERMNO are ignored. Fortunately, these account for a very

small fraction of institutional holdings. Some companies have multiple equity securities associated with them, and CRSP uniquely identifies each firm with a permanent company number, or PERMCO. We value-weighted returns and institutional holding percentages of the different share classes (PERMNOs) associated with each PERMCO. This gives us one return and one institutional ownership percentage associated with each set of accounting data. Our primary results are based on annual vector autoregressions. In these VARs we use end-of-year t-1 accounting information to predict returns from July of year t through June of year t+1: We use institutional ownership data as of June 30 of year t as the variable corresponding to the returns over this period. In some of our tests we use monthly data. For this purpose we compute the percentage of institutional ownership at the end of each quarter and assume that the number stays constant over the subsequent three months. In this way, we can compute monthly returns on the aggregate institutional portfolio. We define all outstanding shares not held by 13F institutions to be ‘‘individual’’ or ‘‘household’’ holdings. (Our individual holdings thus contain assets controlled by very small financial institutions, but these make up only a tiny percentage of the category.) Therefore, for each stock the individual holding fraction is simply one minus the institutional ownership.

6. General Appendix: STATS data

The STATS data produced by analysts of Mogan Stanley Capital Indexes (MSCI) and before 2010 by Kinder Lydenberg and Domini (KLD) is used in three Chapters of this dissertation. STATS provides strengths and concern indicators for Environmental, Social, and Governance (ESG) issues and indicators for controversial business practices and industries with expanding coverage over time starting in 1991 (Table 6.1). In this appendix I list a brief summary of the STATS indicators. And for transparency I provide a thorough description for the environmental indicators. The full description of the data can be downloaded from the WRDS website in the manuals and overviews section.

Table 6.1 Coverage of STATS data

Index Universe	1991-2000	2001	2002	2003-Present
S&P 500 Index	X	X	X	X
Domini 400 Social Index	X	X	X	X
Russell 1000 Index		X	X	X
Large Cap Social Index			X	X
Russell 3000 Index				X
Broad Market Social Index				X
Companies covered	650	1100	1100	3100

-ESG strength and concern indicators

Environmental Strengths

Beneficial Products and Services-The company derives substantial revenues from innovative remediation products, environmental services, or products that promote the efficient use of energy, or it has developed innovative products with environmental benefits. (The term "environmental service" does not include services with questionable environmental effects, such as landfills, incinerators, waste-to-energy plants, and deep injection wells.)

Pollution Prevention-The company has notably strong pollution prevention programs including both emissions reductions and toxic-use reduction programs.

Recycling-The company either is a substantial user of recycled materials as raw materials in its manufacturing processes, or a major factor in the recycling industry.

Clean Energy-The company has taken significant measures to reduce its impact on climate change and air pollution through use of renewable energy and clean fuels or through energy efficiency. The company has

demonstrated a commitment to promoting climate-friendly policies and practices outside its own operations. KLD renamed the Alternative Fuels strength as Clean Energy Strength.

Communications-The company is a signatory to the CERES Principles, publishes a notably substantive environmental report, or has notably effective internal communications systems in place for environmental best practices. KLD began assigning strengths for this issue in 1996, and then incorporated the issue with the Corporate Governance: Transparency rating which was added in 2005. In all spreadsheets it is incorporated into the Transparency rating.

Property, Plant, and Equipment-The company maintains its property, plant, and equipment with above average environmental performance for its industry. KLD has not assigned strengths for this issue since 1995.

Other Strength-The company has demonstrated a superior commitment to management systems, voluntary programs, or other environmentally proactive activities.

Environmental Concerns

Regulatory Compliance-This indicator measures a firm's record of compliance with environmental regulations. Factors affecting this evaluation include, but are not limited to, fines/sanctions for causing environmental damage, and/or violations of operating permits.

Substantial Emissions-This indicator measures a firm's emission of toxic chemicals according to data from the Toxics Release Inventory (TRI), a U.S. Environmental Protection Agency (EPA) database of information on toxic chemical releases and waste management activities. Factors affecting this evaluation include, but are not limited to, how the firm compares to its industry peers.

Climate Change-This indicator measures the severity of controversies related to a firm's climate change related policies and initiatives. Factors affecting this evaluation include, but are not limited to, a history of involvement in greenhouse gas (GHG)-related legal cases, widespread or egregious impacts due to corporate GHG emissions, resistance to improved practices, and criticism by non-governmental organizations (NGOs) and/or other third-party observers. In addition, factors cover whether a company derives substantial revenues from the sale of coal or oil and its derivative fuel products, or whether the company derives substantial revenues indirectly from the combustion of coal or oil and its derivative fuel products.

Negative Impact of Products & Services-This indicator measures the negative environmental impact of a firm's products and/or services. Factors affecting this evaluation include, but are not limited to, products/services that involve regulated substances, the production/consumption of hazardous chemicals, and controversial products such as those that use genetically modified organisms or nanotechnology.

Land Use & Biodiversity-This indicator measures the severity of controversies related to a firm's use or management of natural resources. Factors affecting this evaluation include, but are not limited to, a history of involvement in natural resource-related legal cases, widespread or egregious impacts due to the firm's use of natural resources, resistance to improved practices, and criticism by NGOs and/or other third-party observers.

Non-Carbon Emissions-This indicator measures the severity of controversies related to a firm's non-GHG emissions. Factors affecting this evaluation include, but are not limited to, a history of involvement in land, air, or water emissions-related legal cases, widespread or egregious impacts due to corporate non-GHG emissions, resistance to improved practices, and criticism by NGOs and/or other third-party observers.

Other Concern-This indicator measures the severity of controversies related to a firm's environmental impact. Factors affecting this evaluation include, but are not limited to widespread or egregious environmental impacts, resistance to improved practices, criticism by NGOs and/or other third-party observers, and any other environmental controversies not covered by other environmental ratings.

Community Strengths

Charitable Giving
Innovative Giving
Non-US Charitable Giving
Support for Housing
Support for Education
Indigenous Peoples Relations
Volunteer Programs
Other Strength

Community Concerns

Community Impact
Operations in Burma
Operations in Sudan
Other Concern

Corporate Governance Strengths

Limited Compensation
Ownership Strength
Transparency Strength
Political Accountability Strength
Other Strength

Corporate Governance Concerns

High Compensation
Ownership Concern
Transparency Concern
Polical Accountability Concern
Accounting Concern
Other Concern

Diversity Strengths

Female CEO
Board of Directors
Work/Life Benefits
Women & Minority Contracting
Employment of the Disabled
Gay & Lesbian Policies
Other Strength

Diversity Concerns

Workforce Diversity Controversies
Board of Directors
Representation
Other Concern

Employee Relations Strengths

Union Relations
No-Layoff Policy
Retirement Benefits Strength
Health and Safety Strength
Cash Profit Sharing
Employee Involvement
Other Strength

Employee Relations Concerns

Union Relations
Workforce Reductions
Retirement/Pension Benefits Concern
Health and Safety Concern
Other Concern

Human Rights Strengths	Human Rights Concerns
Positive Record in South Africa	South Africa Controversies
Indigenous Peoples Relations Strength	Indigenous Peoples Relations Concern
Labor Rights Strength	Labor Rights Concern
Other Strength	Northern Ireland Operations
	Burma Concern
	Mexico Concern (labor or environmental)
	Other Concern

N.B. Most of the indicators in the Human Rights area are measured during a small sub period in the data.

Product quality Strengths	Product quality Concerns
Quality Program	Product Safety
R&D/Innovation	Marketing/Contracting Concern
Benefits to Economically Disadvantaged	Antitrust
Other Strength	Customer Relations Concern
	Other

-*Controversial business indicators*

Alcohol

Producer

Distributor

Retailer.

Licensor

Supplier

Ownership of an Alcohol Company

Ownership by an Alcohol Company

Gambling

Operations

Support:

- Products manufactured
- Gambling technology and support
- Gambling-related services

Licensor

Ownership of a Gambling Company

Ownership by a Gambling Company

Tobacco

Producer

Distributor

Retailer.

Licensor

Supplier

Ownership of a Tobacco Company

Ownership by a Tobacco Company

Firearms

Producer

Distributor

Retailer

Ownership of a Firearms Company

Ownership by a Firearms Company

Military

Conventional Weapons Systems & Components

Nuclear Weapons Systems & Components

Chemical and Biological Weapons Systems & Components

Support Systems and Services

Ownership of a Weapons company

Ownership by a Weapons company

Nuclear power

Nuclear Power Generation

Essential Suppliers:

- Builders or designers
- Suppliers of nuclear-specific key components or essential services
- Nuclear industry consulting and license renewal services
- Companies involved in uranium mining, spent fuel processing, and fuel storage
- Distributors and handlers of nuclear fuel
- Repair and maintenance

Ownership of a Nuclear energy company

Ownership by a Nuclear energy company

References

Ali, A., X. Chen, T. Yao and T. Yu. "Do mutual funds profit from the accruals anomaly?", *Journal of Accounting Research*, 46 (2008), 1-26.

Amihud, Y. and R. Goyenko. "Mutual fund's R^2 as predictor of performance." *Review of Financial Studies*, 26 (2013), 667-694.

Andreoni, J., and J. Miller. "Giving according to GARP: An experimental test of the consistency of preferences for altruism." *Econometrica*, 70 (2002), 737-753

Andreoni, J., and L. Vesterlund. "Which is the fair sex? Gender differences in altruism." *Quarterly Journal of Economics*, 116 (2001), 293-312.

Angel, J.J. and P. Rivoli. "Does ethical investing impose a cost upon the firm? A theoretical perspective." *Journal of Investing*, 6 (1997), 57–61.

Auger, P., P. Burke, T.M. Devinney, and J.J. Louviere. "What will consumers pay for social product features?" *Journal of Business Ethics* 42 (2003), 281-304.

Bae, K., J. Kang and J. Wang. "Employee treatment and firm leverage: A test of the stakeholder theory of capital structure." *Journal of Financial Economics*, 100 (2011), 130-253.

Barber, B. "Monitoring the monitor: Evaluating CalPERS' activism." *Journal of Investing* Winter (2007), 66-80.

Barney, J.B. "Firm resources and sustained competitive advantage." *Journal of Management*, 17 (1991), 99–120.

Barsky, R., F.T. Juster, M.S. Kimball, and M.D. Shapiro. "Preference parameters and behavioral heterogeneity: An experimental study in the Health and Retirement Study." *Quarterly Journal of Economics* 112 (1997), 537-579.

Bauer, R., and P. Smeets. "Social identification and investment decisions." *Working Paper* (2010), University of Maastricht, NL.

Bebchuk, L., A. Cohen and C. Wang. "What matters in corporate governance?" Review of Financial Studies, 22 (2009), 783-827.

Bebchuk, L., A. Cohen and C. Wang. "Learning and the disappearing association between governance and stock returns." *Journal of Financial Economics,*108 (2013), 323-348.

Becchetti, L., and F.C. Rosati. "Global social preferences and the demand for socially responsible products: Empirical evidence from a pilot study on fair trade consumers." *The World Economy* 30 (2007), 807-836.

Benartzi, S., and R.H. Thaler. "Naïve diversification strategies in defined contribution savings plans." *American Economic Review* 91 (2005), 79-98.

Benson, K.L., and J.E. Humphrey. "Socially responsible investment funds: Investment reaction to current and past returns." *Journal of Banking and Finance*, 32 (2008), 1850-1859.

Bernheim, D.B., D.M. Garrett, and D.M. Maki. "Education and saving: The long-term effects of high school financial curriculum mandates." *Journal of Public Economics* 80 (2001), 435-465.

Bernheim, D.B., D.M. Garrett. "The effects of financial education in the workplace: evidence from a survey of households." *Journal of Public Economics* 87 (2003), 1487-1519.

Bessire, D. and S. Onnée. “Assessing corporate social performance: Strategies of legitimation and conflicting ideologies.” *Critical Perspectives on Accounting*, 21 (2010), 445–467.

Bollen, N. “Mutual fund attributes and investor behavior.” *Journal of Financial and Quantitative Analysis*, 42 (2007), 683-708.

Borgers, A., J. Derwall, K. Koedijk, and J. Ter Horst. “Stakeholder relations and stock returns: On errors in investors’ expectations and learning.” *Journal of Empirical Finance* 22 (2013), 159-175.

Borgers, A., J. Derwall, K. Koedijk, and J. Ter Horst. “Can investors profit from social tastes? Evidence from mutual fund holdings.” *Working paper* (2014).

Borgers, A. and R. Pownall. “Attitudes towards environmentally and socially responsible investment.” *Journal of Behavioral and Experimental Finance*, 1 (2014), 22-44.

Calvet, L.E., J.Y. Campbell, and P. Sodini. “Measuring the financial sophistication of households.” *American Economic Review*, 99 (2009a), 393-398.

Calvet, Laurent E., John Y. Campbell, and Paulo Sodini. “Fight or flight? Portfolio rebalancing of individual investors.” *Quarterly Journal of Economics*, 124 (2009b), 301-348.

Campbell, J.Y. “A variance decompositions for stock returns.” *Economic Journal* 101, 57-179.

Campbell, J.Y., C. Polk, and T. Vuolteenaho (2010), “Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns”, Review of Financial Studies 23 (2010), 305-344.

Carhart, M. “On the persistence in mutual fund performance.” *Journal of Finance,* 52 (1997), 57–82.

Chan, L., N. Jegadeesh and J. Lakonishok. “Momentum strategies.” *Journal of Finance, 51* (1996), 1681-1713.

Chava, S. “Environmental externalities and the cost of capital.” (2013) *Management Science,* forthcoming.

Chen, I.H., H. Hong, and K. Shue. “Do managers do good with other people’s money?” *Working Paper* (2013), Princeton University.

Chordia, T., Subrahmanyam, A. and Q. Tong. “Trends in the Cross-Section of Expected Stock Returns”, *Working Paper* (2012), University of California at Los Angeles.

Cohen, R. B., P.A. Gompers and T. Vuolteenaho (2002), ”Who underreacts to cash-flow news? Evidence from trading between individuals and institutions”, *Journal of Financial Economics* , 66, 409-462.

Core, J. R. W. Guay, and T. Rusticus. “Does weak governance cause weak stock returns? An examination of firm operating performance and investors’ expectations.” *Journal of Finance,* 61 (2006), 655–687.

Cornell, B. and A. Shapiro. “Corporate stakeholders and corporate finance.” *Financial Management,* 16 (1987), 5-14.

Cremers, M., Petajisto, A. and E. Zitzewitz. “Should benchmark indices have alpha? Revisiting performance evaluation. (2010), *Critical Finance Review*, forthcoming.

Dahliwal, D., O. Li, A. Tsang and Y. Yang. “Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting.” *The Accounting Review, 86* (2011), 59-100.

DeBondt, W., and R. Thaler. “Does the stock market overreact? *Journal of Finance* 40 (1985), 557-581.

Derwall, J., N. Guenster, R. Bauer and K. Koedijk. “The eco-efficiency premium puzzle.” *Financial Analysts Journal,* 61 (2005), 51–63.

Deng, X., J. Kang, and B.S. Low. Corporate social responsibility and sakeholder value maximization: Evidence from mergers.” *Journal of Financial Economics* (2013) 110, 87-109.

Derwall, J., K. Koedijk and J. Ter Horst. “A tale of values-driven and profit-seeking social investors.” *Journal of Banking and Finance,* 35 (2011), 2137-2147.

Di Giuli, A. and L. Kostovetsky. “Are red or blue companies more likely to go green? Politics and coporate social responsibility.” *Journal of Financial Economics* (2013), Forthcoming.

Dimson, E., O. Karakas, and X. Li. “Active ownership” *Working Paper* (2012), LBS.

Doukas, J., C. Kim, and C. Pantzalis. “A Test of the errors–in–expectations explanation of the value/glamour stock returns performance: evidence from analysts’ forecasts.” *Journal of Finance,* 57 (2002), 2143-2165.

Dowell, G., S. Hart and B. Yeung. “Do corporate global environmental standards create or destroy market value?” *Management Science, 46* (2000), 1059-1074.

Durand, R.B., S. Koh, and P.L. Tan. “The price of sin in the Pacific-Basin.” *Pacific-Basin Finance Journal*, 21 (2013) 899-913.

Edmans, A. “Does the stock market fully value intangibles? Employee satisfaction and equity prices.” *Journal of Financial Economics*, 101 (2011), 621-640.

El Ghoul, S. O. Guedhami, C.C.Y. Kwok and D.R. Mishra. “Does corporate social responsibility affect the cost of capital?" *Journal of Banking and Finance*, 35 (2011), 2388-2406.

Els, P.J.A. Van, J.W. Van Den End, M.C.J. Van Rooij. “Pensions and public opinion: a survey among Dutch households.” *De Economist* 152 (2004), 101-116.

Fabozzi, F. J., K.C. Ma, B.J. Oliphant. “Sin stock returns”, *Journal of Portfolio Management*, 35 (2008), 82–94.

Fama, E.F. and K.R. French. “Common risk factors in the returns on stocks and bonds*.”, Journal of Financial Economics*, 33 (1993) 3-56.

Fama, E.F. and K.R. French. “Disagreement, tastes, and asset pricing.” *Journal of Financial Economics*, 83 (2007), 667–689.

Fauver, L. and M.B. McDonald. “Interntional variation in sin stocks and its effects on equity valuation.” *Journal of Corporate Finance,* 25 (2014), 173-187.

Galema, R., A. Plantinga and B. Scholtens. “The stocks at stake: Return and risk in socially responsible investing.” *Journal of Banking and Finance***,** 32 (2008), 2646–2654.

Gebhardt, W.R., C.M.C. Lee, and B. Swaminathan. “Toward an implied cost of capital.” *Journal of accounting research,* 39 (2001), 135-176.

Gollier, C. and S. Pouget. “Asset prices and corporate behavior with socially responsible investors.” Working Paper, Toulouse School of Economics (2012).

Gompers, P., J. Ishii and A. Metrick. “Corporate governance and equity prices.” *Quarterly Journal of Economics,* 118 (2003), 107–155.

Guiso, L., P. Sapienza, and L. Zingales. “Trusting the stock market.” *Journal of Finance* 63 (2008), 2557-2600.

Hart, S. “A natural-resource-based view of the firm.” *The Academy of Management Review, 20 (*1995), 986-1014.

Heinkel, R., A. Kraus and J. Zechner. “The effect of green investment on corporate behavior.” *Journal of Financial and Quantitative Analysis*, 35 (2001), 431–449.

Hong, H. and M. Kacperczyk. “The price of sin: The effects of social norms on markets”. *Journal of Financial Economics*, 93 (2009), 15-36.

Hong, H. and L. Kostovetsky. “Red and blue investing: Values and finance”. *Journal of Financial Economics,* 103 (2012), 1-19.

Huij, J. and M. Verbeek. “On the use of multifactor models to evaluate mutual fund performance.” *Financial Management*, 38 (2009), 75-102.

Jiao, Y. “Stakeholder welfare and firm value.” *Journal of Banking and Finance,* 34 (2010), 2549-2561.

Kapetanios, G. (2008). “A bootstrap procedure for panel data sets with many cross-sectional units.” *The Econometrics Journal*, 11, 377-395.

Kempf, A. and P. Osthoff. “The effect of socially responsible investing on portfolio performance.” *European Financial Management*, 13 (2007), 908-922.

Kempf, A. and P. Osthoff. “SRI funds: Nome nest omen.” *Journal of Business Finance and Accounting*, 35 (2008), 1276-1294.

Kim, I. and M. Venkatachalam. “Are sin stocks paying the price for accounting sins?” *Journal of Accounting, Auditing and Finance,* 26 (2011), 415-422.

King, A. and M. Lenox. “Exploring the locus of profitable pollution reduction.” *Management Science*, 48 (2002), 289–99.

Kumar, A. and J. Page. “Deviations From Norms and Informed Trading,” *Journal of Financial and Quantitative Analysis,* Forthcoming, 2011.

Kumar, A., J. Page, and O. Spalt. “Religious beliefs, gambling attitudes, and financial market outcomes.” *Journal of Financial Economics* 102 (2011), 671-708.

La Porta, R., J. Lakonishok, A. Shleifer and R.W. Vishny, “Good news for value stocks: further evidence on market efficiency.” *Journal of Finance,* 50 (1997), 1715–1742.

Laroche, M., J. Bergeron, and G. Barbaro-Forleo. “Targeting consumers who are willing to pay more for environmentally friendly products.” *Journal of Consumer Marketing,* 18 (2001), 503-520.

Lim, T. “Rationality and analysts’ forecast bias.” *Journal of Finance,* 56 (2001), 369-385.

Liston, D.P. and G. Soydemir. “Faith-based and sin portfolios: An empirical inquiry into norm-neglect vs. norm-conforming investor behavior.” *Managerial Finance,* 36 (2010), 876-885.

Ljungqvist, A., C. Malloy, and F. Marston. “Rewriting history.” *Journal of Finance,* 64 (2009), 1935-1960.

Lo, A.W. and A.C. MacKinlay. "Data-snooping biases in tests of financial asset pricing models." *Review of Financial Studies*, 3 (1990), 431-67.

Lobe, S. and C. Walkhausl. “Vice vs. virtue investing around the world.” *Working Paper University of Regensburg (2011).*

Lusardi, A., and O.S. Mitchell. “Baby Boomer Retirement Security: The Role of Planning, Financial Literacy, and Housing Wealth.” *Journal of Monetary Economics* 54 (2007a), 205-224.

Lusardi, A., and O.S. Mitchell. “Financial literacy and retirement preparedness: Evidence and implications for financial education.” *Business Economics* 42 (2007b), 35-44.

Lusardi, A., and O.S. Mitchell. “Planning and Financial Literacy: How do women Fare?” *American Economic Review* 98 (2008), 413-417.

McWilliams, A., and D. Siegel. “Corporate social responsibility and financial performance: Correlation or misspecification?” *Strategic Management Journal* 21 (2000), 603-609.

Morse, A., and S. Shive. “Patriotism in your portfolio.” *Journal of Financial Markets,* 14 (2011), 411-440.

Nevins, J.L., W.O. Bearden, and B. Money, "Ethical values and long-term orientation," *Journal of Business Ethics*, 2007, *71* (3), 261–274.

Newey, W.K. and K.D. West. "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica*, 55 (1987), 703-708.

Nilsson, J. "Segmenting socially responsible mutual fund investors: the influence of financial return and social responsibility." *International Journal of Bank Marketing* 27 (2009), 5-31.

O'Loughlin, J. and R. Thamotheram, "Enhanced Analytics for a New Generation of Investor", Report by Universities Superannuation Scheme Ltd, (2006).

Pastor, L. and R.F. Stambaugh, "Liquidity risk and expected stock returns," *Journal of Political Economy*, vol. 111 (2003), 642-685.

Quandt, R.E. "Tests of the hypothesis that a linear regression system obeys two separate regimes." *Journal of the American Statistical Association*, 55 (1960), 324-330.

Renneboog, L., and C. Spaenjers. "Religion, economic attitudes, and household finance." *Oxford Economic Papers* 64 (2011), 103-127.

Renneboog, L., J. Ter Horst, and C. Zhang. "Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds." *Journal of Financial Intermediation* 20 (2011), 562-588.

Richardson, B. "Do the fiduciary duties of pension funds hinder socially responsible investment?" *Banking and Finance Law Review* 22 (2007), 146-201.

Richardson, B. "Keeping ethical investment ethical: regulatory issues for investing for sustainability." *Journal of Business Ethics* 87 (2009), 555-572.

Richardson, B. "From fiduciary duties to fiduciary relationships for socially responsible investing: responding to the will of beneficiaries." *Journal of Sustainable Finance and Investment* 1 (2011), 5-19.

Rooij, M.C.J. Van, C.J.M. Kool, and H.M. Prast. "Risk-return preferences in the pension domain: Are people able to choose?" *Journal of Public Economics* 91 (2007), 701-722.

Rooij, M.C.J. Van, A. Lusardi, and R.J.M. Alessie. "Financial literacy and retirement planning in the Netherlands." *Journal of Economic Psychology* 32 (2011a), 593-608.

Rooij, M.C.J. Van, A. Lusardi, and R.J.M. Alessie. "Financial literacy and stock market participation." *Journal of Financial Economics* 101 (2011b), 449-472.

Russo, M. and P. Fouts. "A resource-based perspective on corporate environmental performance and profitability." *Academy of Management Journal*, 40 (1997), 534–59.

Salaber, J. "The determinants of sin-stock returns: Evidence on the European market." *Working Paper University of Bath (2007).*

Schuyt T.N.M., B.M. Gouwenberg, R.H.F.P. Bekkers. "Geven in Nederland 2009; Giften, Nalatenschappen, Sponsoring en Vrijwilligerswerk." Reed Business, (2009), 57-64.

Schwert, G. "Anomalies and market efficiency." In G.M. Constantinides & M. Harris & R. M. Stulz (ed.), *Handbook of the Economics of Finance, ed. 1*, Elsevier, (2003), 939-974.

Shumway, T., (1997). "The delisting bias in CRSP data." *Journal of Finance* 41, 579-590.

Sirri, E.R. and P. Tufano, "Costly search and mutual fund flows." *Journal of Finance*, 53 (1998), 1589-1622.

Sloan, R. "Do stock prices fully reflect information in accruals and cash flows about future earnings?" *The Accounting Review*, 71 (1996), 280-315.

Statman, M. and D. Glushkov. “The wages of social responsibility.” *Financial Analysts Journal,* 65 (2009), 33–46.

Verwijmeren, P. and J. Derwall. “Employee well-being, firm leverage, and bankruptcy risk.” *Journal of Banking and Finance,* 34 (2010), 956-964.

Vuolteenaho, T. (2002). “What drives firm-level stock returns?” *The Journal of Finance*, 57, 233-264.

Wernerfelt, B. “The resource-based view of the firm.” *Strategic Management Journal*, 5 (1984), 171–180.

Zingales, L. “In search of new foundations.” *Journal of Finance,* 55 (2000), 1623-1653.